

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:14:39 ; Search time 229.28 Seconds
(without alignments)
1.995 Million cell updates/sec

Title: PCT-US01-05625A-20

Perfect score: 41

Sequence: 1 GGLLVQDC 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_0401.*
2: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
23: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	8	21	Y79124
2	38	92.7	8	21	Y79120
3	36	87.8	8	21	Y79122
4	34	82.9	8	21	Y79123
5	33	80.5	8	21	Y79112
6	33	80.5	8	21	Y79118
7	33	80.5	8	21	Y79128
8	33	80.5	115	20	Y42459
9	33	80.5	115	20	Y42460
10	33	80.5	362	21	Y82497
11	33	80.5	377	21	Y68643

12	33	80.5	380	21	Y68659	Amino acid sequenc
13	33	80.5	380	21	Y68662	Amino acid sequenc
14	32	78.0	115	20	Y27892	Human secreted pro
15	32	78.0	116	21	Y91324	Group B Streptococ
16	32	78.0	162	20	Y59651	Secreted protein e
17	32	78.0	162	20	Y25457	Human secreted pro
18	32	78.0	162	20	Y36051	Extended human sec
19	32	78.0	162	20	Y35888	Extended human sec
20	32	78.0	162	20	W93619	Human 5' EST seque
21	32	78.0	162	20	Y12678	Human 5' EST seque
22	32	78.0	162	20	Y12985	Human secreted pro
23	32	78.0	162	20	Y04172	Human 5' EST seque
24	32	78.0	162	20	Y01593	Secreted protein e
25	32	78.0	162	20	Y04154	Human 5' EST seque
26	31	75.6	8	21	Y84661	Peptide antagonist
27	31	75.6	8	21	Y79119	Human foetal intes
28	31	75.6	20	20	W94489	Human foetal intes
29	31	75.6	20	21	Y79133	Human foetal intes
30	31	75.6	25	15	R52775	Murine BRE-3 immun
31	31	75.6	25	15	R52793	Murine BRE-3 immun
32	31	75.6	92	21	B54096	Human pancreatic c
33	31	75.6	92	21	G49732	Arabidopsis thalia
34	31	75.6	130	21	G27097	Zea mays protein f
35	31	75.6	164	21	G27095	Zea mays protein f
36	31	75.6	166	21	G49731	Arabidopsis thalia
37	31	75.6	196	20	Y36831	Amino acid sequenc
38	31	75.6	295	21	Y75121	Neisseria meningit
39	31	75.6	295	21	Y75122	Neisseria meningit
40	30	73.2	8	21	Y79108	Peptide antagonist
41	30	73.2	196	19	Y86035	S. pneumoniae derl
42	30	73.2	355	20	Y32173	Chlamydia psittaci
43	30	73.2	368	16	R67970	Alfalfa beta-1,3-g
44	30	73.2	492	20	W98121	Arabidopsis endo-1
45	30	73.2	498	21	B07577	Protein encoded by

ALIGNMENTS

RESULT 1	Y79124 standard; Peptide; 8 AA.
ID	Y79124
XX	Y79124;
AC	05-JUN-2000 (first entry)
DT	XX
XX	Peptide antagonist of zonulin.
DE	XX
KW	Zonulin; antagonist; zonula occludens toxin receptor;
KW	blood-brain barrier; antiinflammatory; cerebroprotective;
KW	neuroprotective; dermatological; antitumor; antiviral;
KW	antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
KW	hypotensive; immunosuppressive; antiparasitic; vasotropic;
KW	gastrointestinal inflammation; therapy.
XX	XX
OS	Synthetic.
XX	XX
PN	W0200007609-A1.
XX	XX
PD	17-FEB-2000.
XX	XX
PF	28-JUL-1999; 99WO-US16683.
XX	XX
PR	03-AUG-1998; 98US-0127815.
XX	XX
PA	(UYMA-) UNIV MARYLAND BALTIMORE.
XX	XX
PI	Fasano A;
XX	XX
DR	WPI; 2000-205565/18.
XX	XX
PT	New peptide antagonist of zonulin useful as antiinflammatory agent for

PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 XX
 PS Claim 1: Page 47: 69pp: English.
 CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 CC
 SQ Sequence 8 AA:
 QY 1 GGLVNDG 8
 DB 1 99LVVQDQ 8
 Query Match 100.0%; Score 41; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Y79120 standard; Peptide; 8 AA.
 Y79120;
 05-JUN-2000 (first entry)
 Peptide antagonist of zonulin.
 Zonulin; antagonist; zonula occludens toxin receptor;
 blood-brain barrier; antiinflammatory; cerebroprotective;
 neuroprotective; dermatological; antitumor; antiviral;
 antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
 hypotensive; immunosuppressive; antiparasitic; vasotropic;
 gastrointestinal inflammation; therapy.
 Synthetic.
 WO200007609-A1.
 17-FEB-2000.
 28-JUL-1999; 99WO-US16683.
 03-AUG-1998; 98US-0127815.
 (UYMA-) UNIV MARYLAND BALTIMORE.

PI Fasano A;
 XX WPI: 2000-205565/18.
 DR
 XX
 PT New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 XX
 PS Claim 1: Page 46: 69pp: English.
 CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 CC
 SQ Sequence 8 AA:
 QY 1 GGLVNDG 8
 DB 1 99LVVQDQ 8
 Query Match 92.7%; Score 38; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Y79122 standard; Peptide; 8 AA.
 Y79122;
 05-JUN-2000 (first entry)
 Peptide antagonist of zonulin.
 Zonulin; antagonist; zonula occludens toxin receptor;
 blood-brain barrier; antiinflammatory; cerebroprotective;
 neuroprotective; dermatological; antitumor; antiviral;
 antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
 hypotensive; immunosuppressive; antiparasitic; vasotropic;
 gastrointestinal inflammation; therapy.
 Synthetic.
 WO200007609-A1.
 17-FEB-2000.
 28-JUL-1999; 99WO-US16683.

XX 03-AUG-1998: 98US-0127815.
 PR (UYMA-) UNIV MARYLAND BALTIMORE.
 PA
 XX Fasano A;
 PI
 DR WPI: 2000-205565/18.
 XX
 PT New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -
 XX
 PS Claim 1: Page 46; 69pp: English.
 XX
 CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulin, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 CC
 XX
 XX Sequence 8 AA:
 SQ
 Query Match 87.8%: Score 36; DB 21; Length 8;
 Best Local Similarity 87.5%: Pred. No. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 GGLVODG 8
 Db 1 gglevqdg 8
 RESULT 4
 Y79123
 ID Y79123 standard: Peptide: 8 AA.
 AC Y79123:
 DT 05-JUN-2000 (first entry)
 DE Peptide antagonist of zonulin.
 XX
 KM Zonulin; antagonist; zonula occludens toxin receptor;
 KM blood-brain barrier; antiinflammatory; cerebroprotective;
 KM neuroprotective; dermatological; antiulcer; antiviral;
 KM antibacterial; cytostatic; anti-HIV; vulnerary; antiallergic;
 KM hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KM gastrointestinal inflammation; therapy.
 XX
 XX Synthetic.
 OS
 XX

PN W0200007609-A1.
 XX
 XX 17-FEB-2000.
 PD
 XX 28-JUL-1999: 99WO-US16683.
 PF
 XX
 PR 03-AUG-1998: 98US-0127815.
 XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 PI
 DR WPI: 2000-205565/18.
 XX
 PT New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -
 XX
 PS Claim 1: Page 47; 69pp: English.
 XX
 CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulin, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 CC
 XX
 XX Sequence 8 AA:
 SQ
 Query Match 82.9%: Score 34; DB 21; Length 8;
 Best Local Similarity 87.5%: Pred. No. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 GGLVODG 8
 Db 1 gglivqdg 8
 RESULT 5
 Y79112
 ID Y79112 standard: Peptide: 8 AA.
 AC Y79112:
 DT 05-JUN-2000 (first entry)
 DE Peptide antagonist of zonulin.
 XX
 KM Zonulin; antagonist; zonula occludens toxin receptor;
 KM blood-brain barrier; antiinflammatory; cerebroprotective;
 KM neuroprotective; dermatological; antiulcer; antiviral;
 KM antibacterial; cytostatic; anti-HIV; vulnerary; antiallergic;
 KM

KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.
 OS Synthetic.
 KM WO200007609-A1.
 KM 17-FEB-2000.
 KM 28-JUL-1999; 99WO-US16683.
 KM 03-AUG-1998; 98US-0127815.
 KM (UYMA-) UNIV MARYLAND BALTIMORE.
 KM Fasano A;
 KM WPI; 2000-205565/18.
 DR New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 XX Claim 1; Page 43; 69pp; English.
 PS
 XX This present sequence is that of a peptide antagonist of zonulin
 CC (1), one of 25 such peptides (see 179105-23) of the invention,
 CC which bind to a zonulin receptor (ZOR) receptor, yet do not
 CC physiologically modulate the opening of zonulin tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOR
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOR receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC celiac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 CC
 XX Sequence 8 AA:
 SQ
 Query Match 80.5%; Score 33; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGLVQDG 8
 Db 1 grllvqdg 8
 RESULT 6
 ID Y79118 standard; Peptide; 8 AA.
 AC Y79118;
 KM 05-JUN-2000 (first entry)
 KM Peptide antagonist of zonulin.
 DE

XX Zonulin; antagonist; zonula occludens toxin receptor;
 KM blood-brain barrier; antiinflammatory; cerebroprotective;
 KM neuroprotective; dermatological; antileuk; antiviral;
 KM antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
 KM hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KM gastrointestinal inflammation; therapy.
 OS Synthetic.
 KM WO200007609-A1.
 KM 17-FEB-2000.
 KM 28-JUL-1999; 99WO-US16683.
 KM 03-AUG-1998; 98US-0127815.
 KM (UYMA-) UNIV MARYLAND BALTIMORE.
 KM Fasano A;
 KM WPI; 2000-205565/18.
 DR New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 XX Claim 1; Page 45; 69pp; English.
 PS
 XX This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see 179105-23) of the invention,
 CC which bind to a zonulin receptor (ZOR) receptor, yet do not
 CC physiologically modulate the opening of zonulin tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOR
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOR receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC celiac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 CC
 XX Sequence 8 AA:
 SQ
 Query Match 80.5%; Score 33; DB 21; Length 8;
 Best Local Similarity 75.0%; Pred. No. 3.2e+05;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGLVQDG 8
 Db 1 gycvqdg 8
 RESULT 7
 ID Y79128 standard; Peptide; 8 AA
 XX

AC Y79128;
 XX
 DT 05-JUN-2000 (first entry)
 XX
 DE Peptide antagonist of zonulin.
 XX
 KM Zonulin: antagonist; zonula occludens toxin receptor;
 KM blood-brain barrier; antiinflammatory; cerebroprotective;
 KM neutroprotective; dermatological; antiviral;
 KM antibacterial; cytostatic; anti-HIV; vulnerrary; antiallergic;
 KM hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KM gastrointestinal inflammation; therapy.
 XX
 OS Synthetic.
 XX
 PN WO200007609-A1.
 XX
 PD 17-FEB-2000.
 XX
 PF 28-JUL-1999; 99WO-US16683.
 XX
 PR 03-AUG-1998; 98US-0127815.
 XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Pasano A;
 XX
 DR WPI: 2000-205565/18.
 XX
 PT New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -
 XX
 PS Claim 1: Page 48; 69pp; English.
 XX
 CC This present sequence is that of a peptide antagonist of zonulin
 CC (2). One of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infection, bacterial overgrowth, whipple's disease, diseases with
 CC collagen erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 XX
 SO Sequence 8 AA:
 Query Match 80.5%; Score 33; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 8
 ID Y42459 standard; Protein: 115 AA.
 XX
 AC Y42459;
 XX
 DT 21-DEC-1999 (first entry)
 XX
 DE Human guanylin preprohormone sequence.
 XX
 KM Human; guanylin; preprohormone; intestinal guanylate cyclase regulator;
 KM bacterium; heat stable enterotoxin; laxative; constipation.
 XX
 OS Homo sapiens.
 XX
 PN US5969097-A.
 XX
 PD 19-OCT-1999.
 XX
 PF 23-JUN-1992; 92US-0903029.
 XX
 PR 23-JUN-1992; 92US-0903029.
 XX
 PA (SEAR) SEARLE & CO G. D.
 XX
 PI Wiegand RC, Currie MG, Fok KF;
 XX
 DR WPI: 1999-590442/50.
 DR N-PSDB; 222677.
 XX
 PT Isolated protein used as a laxative in the treatment of constipation -
 XX
 PS Disclosure: Fig 1; 14pp; English.
 XX
 CC This sequence represents a human guanylin preprohormone which is
 CC cleaved to generate a 15 amino acid mature peptide (Y42458). This
 CC sequence is the encoded protein from a composite DNA sequence isolated
 CC by probing a human duodenum cDNA library with the corresponding rat
 CC coding sequence. Guanylin is an intestinal guanylate cyclase regulator
 CC which has similarity to bacterial heat stable enterotoxins which
 CC activate intestinal guanylate cyclases. The guanylin peptide can be
 CC used as a laxative in the treatment of constipation.
 XX
 SO Sequence 115 AA:
 Query Match 80.5%; Score 33; DB 20; Length 115;
 Best Local Similarity 75.0%; Pred. No. 37;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGLVQDG 8
 II: IIII
 Db 20 ggvtvgdg 27

RESULT 9
 ID Y42460 standard; Protein: 115 AA.
 XX
 AC Y42460;
 XX
 DT 21-DEC-1999 (first entry)
 XX
 DE Human guanylin preprohormone sequence.
 XX
 KM Human; guanylin; preprohormone; intestinal guanylate cyclase regulator;
 KM bacterium; heat stable enterotoxin; laxative; constipation.

XX OS Homo sapiens.
 XX PN US5969097-A.
 XX PD 19-OCT-1999.
 XX PE 23-JUN-1992; 92US-0903029.
 XX PR 23-JUN-1992; 92US-0903029.
 XX PA (SEAR) SEARLE & CO G D.
 XX PI Wiegand RC, Currie MG, Fok KF;
 XX WPI: 1999-590442/50.
 XX DR
 XX PT Isolated protein used as a laxative in the treatment of constipation -
 XX PS Disclosure; Fig 2; 14pp; English.
 CC This sequence represents a human guanylin preprohormone which is cleaved
 CC to generate a 15 amino acid mature peptide (Y42438). Guanylin is an
 CC intestinal guanylate cyclase regulator which has similarity to bacterial
 CC heat stable enterotoxins which activate intestinal guanylate cyclases.
 CC The guanylin peptide can be used as a laxative in the treatment of
 CC constipation.
 CC
 SO Sequence 115 AA:

Query Match 80.5%; Score 33; DB 20; Length 115;
 Best Local Similarity 75.0%; Pred. No. 37;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGLVQDG 8
 ||: |||
 Db 20 gsvlcvgd 27

RESULT 10
 Y82497 ID Y82497 standard; protein; 362 AA.
 XX AC Y82497;
 XX DT 12-JUL-2000 (first entry)
 XX DE Human adenovirus type 9 fibre knob protein sequence SEQ ID NO:3.
 XX KW Adenovirus; adenoviral; Ad; fibre knob; fibre protein; gene therapy;
 XX KM Infection; VITRO.
 XX OS Human adenovirus type 9.
 XX PA WO200015823-A1.
 XX PD 23-MAR-2000.
 XX PE 10-SEP-1999; 99WO-US20728.
 XX PR 11-SEP-1998; 98US-009851.
 XX PR 28-MAY-1999; 99US-0136529.
 XX PA (GENV-) GENVEC INC.
 XX PI Michham TJ, Kovesdi I, Roelvink PW, Bruder JT;
 XX WPI: 2000-271459/23.
 PT Recombinant fiber protein used for creating vectors for delivering a
 PT desired gene to a cell with minimal ectopic infection, comprises an
 PT amino terminus of an adenoviral fiber protein and a trimerization

PT domain -
 XX Disclosure; Page 45-46; 69pp; English.
 XX
 CC The present invention describes a recombinant fibre protein (1)
 CC comprising an amino terminus of an adenoviral fibre protein and a
 CC trimerisation domain. The trimerisation domain comprises (a) an
 CC adenoviral fibre knob domain having a mutation affecting at least one
 CC amino acid residue within the region corresponding to the AB loop,
 CC B sheet, DE loop or FG loop of the wild-type adenovirus (Ad) 5 fibre
 CC protein, and where the recombinant fibre protein trimerises when
 CC produced in a eukaryotic cell; or (b) an adenoviral fibre knob domain
 CC having a mutation affecting at least one amino acid residue
 CC corresponding to residue positions 404-406, 408, 409, 412-417, 420, 439,
 CC 441, 442, 445-454, 458, 460, 462, 466, 467, 469-472, 474-477, 482, 485,
 CC 487-492, 505-512, 515, 517, 519, 521-528, 533, 535, 537-549, 551, 553,
 CC 555, 559-568, 580 or 581 of the wild-type Ad5 fibre protein, and where
 CC the recombinant fibre protein trimerises when produced in a eukaryotic
 CC cell. (1) are used for creating vectors which are useful in a variety
 CC of gene-transfer applications, in vitro and in vivo, for delivering a
 CC desired gene to a cell with minimal ectopic infection. Particularly,
 CC the vectors are useful for infecting a desired cell type, an approach
 CC known as alternative targeting. (1) can be used in receptor-ligand
 CC assays and as adhesion proteins in vitro or in vivo. The recombinant
 CC fibre protein permits more efficient production and construction of
 CC safer vectors for gene transfer applications. The present sequence
 CC represents the human adenovirus type 9 fibre knob protein sequence,
 CC which is used in the exemplification of the present invention.
 CC
 SO Sequence 362 AA:

Query Match 80.5%; Score 33; DB 21; Length 362;
 Best Local Similarity 75.0%; Pred. No. 12e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGLVQDG 8
 ||: |||
 Db 67 ggltlqdg 74

RESULT 11
 Y68643 ID Y68643 standard; Protein; 377 AA.
 XX AC Y68643;
 XX DT 05-MAY-2000 (first entry)
 XX DE Amino acid sequence of the fiber protein of Adenovirus serotype 9.
 XX KW Chimeric adenovirus; gene therapy; antigenicity; fiber protein;
 XX KM serotype 9; penton protein; hexon protein.
 XX OS Adenovirus.
 XX PA Key Location/Qualifiers
 XX FT Misc-difference 1..29 /note- "part of the tail of adenovirus serotype 5"
 XX FT
 XX PD WO200003029-A2.
 XX PE 20-JUN-2000.
 XX PR 08-JUL-1999; 99WO-NL00436.
 XX PR 08-JUL-1998; 98EP-0202297.
 XX PA (INTNR-) INTROGENE BV.
 XX PI Havenga M, Vogels R, Bout A;
 XX WPI: 2000-171149/15

```
xx New chimaeric adenoviruses containing a genome derived from different
pt adenovirus serotypes, useful in gene therapy -
ps Example 2; Fig 7; 92pp: English.
xx
CC Y68642-70 represent the amino acid sequences of the fiber proteins
CC of Adenovirus serotypes B, 9, 13, 14, 20, 23, 24, 25, 27, 28, 29, 30,
CC 32, 33, 34, 35, 36, 37, 38, 39, 42, 43, 44, 45, 46, 47, 48, 49 and 51.
CC The proteins are used in the course of the invention to construct
CC chimaeric adenoviruses with reduced antigenicity. The chimaeric
CC adenoviruses comprise at least part of a fiber protein of an adenovirus
CC serotype providing the chimaeric virus with a desired host range and at
CC least part of a penton or hexon protein from another, less antigenic,
CC serotype. The chimaeric adenoviruses are useful for gene therapy,
CC especially where repeated delivery is required. Adenoviruses of the
CC invention are useful can be constructed to have a desired host range and
CC a diminished capability to raise neutralizing antibodies, an absence of,
CC decreased infection of, antigen presenting cells of the immune system
CC (e.g. macrophages), and an ability to escape trapping in the liver
CC through increased target cell specificity.
SO Sequence      377 AA:

Query Match          80.5%; Score 33; DB 21; Length 377;
Best Local Similarity    75.0%; Pred. No. 1.3e+02;
Matches   6; Conservative   1; Mismatches     1; Indels       0; Gaps        0;

OY         1 GGILVQDG 8
           ||| :|||
Db          82 gqlltldg 89

RESULT  12
ID       Y68659 standard; Protein; 380 AA.
XX AC
XX Y68659;
DT 05-MAY-2000 (first entry)
DE Amino acid sequence of the fiber protein of Adenovirus serotype 37.
XX KW Chimaeric adenovirus; gene therapy; antigenicity; fiber protein;
KW serotype 37; penton protein; hexon protein.
XX OS Adenovirus.
XX Key Location/Qualifiers
FH Misc-difference 1..29
FT /note= "part of the tail of adenovirus serotype 5"
XX WO200003029-A2.
PD 20-JAN-2000.
XX PF 08-JUL-1999; 99WO-NL00436.
XX PR 08-JUL-1998; 98EP-O202297.
XX PA (INTR-) INTROGENE BV.
PI Havenga M, Vogels R, Bout A;
DR WPT: 2000-171149/15.
XX PT New chimaeric adenoviruses containing a genome derived from different
XX adenovirus serotypes, useful in gene therapy -
PS Example 2; Fig 7; 92pp: English.
XX Y68642-70 represent the amino acid sequences of the fiber proteins
```

```
CC CC Adenovirus serotypes 8, 9, 13, 14, 20, 23, 24, 25, 27, 28, 29, 30,  
CC 32, 33, 34, 35, 36, 37, 38, 39, 42, 43, 44, 45, 46, 47, 48, 49 and 51.  
CC The proteins are used in the course of the invention to construct  
CC chimaeric adenoviruses with reduced antigenicity. The chimaeric  
CC adenoviruses comprise at least part of a fiber protein of an adenovirus  
CC serotype providing the chimaeric virus with a desired host range and at  
CC least part of a penton or hexon protein from another, less antigenic,  
CC serotype. The chimaeric adenoviruses are useful for gene therapy,  
CC especially where repeated delivery is required. Adenoviruses of the  
CC invention are useful can be constructed to have a desired host range and  
CC a diminished capability to raise neutralizing antibodies, an absence of,  
CC or decreased infection of, antigen presenting cells of the immune system  
CC (e.g. macrophages), and an ability to escape trapping in the liver  
CC through increased target cell specificity.
```

Sequence 380 AA;

```
Query Match            80.5%;     Score 33;   DB 21;   Length 380;  
Best Local Similarity    75.0%;     Pred. No. 1.3e+02;  
Matches    6;   Conservative       1;   Mismatches    1;   Indels       0;   Gaps       0;
```

QY 1 GGLIWODG 8
 ||| :|||
Db 82 ggltltdg 89

RESULT 13

V68662 ID V68662 standard; Protein: 380 AA.

AC Y68662;

XX DT 05-MAY-2000 (first entry)

DE Amino acid sequence of the fiber protein of Adenovirus serotype 42.

XX XX Chimaeric adenovirus; gene therapy; antigenicity; fiber protein;

KW serotype 42; penton protein; hexon protein.

XX OS Adenovirus.

XX FH Key Location/Qualifiers

FT Misc-difference 1..31 /note= "part of the tail of adenovirus serotype 5"

FT FT Misc-difference 237 /note= "not specified"

PN WO200003029-A2.

XX PD 20-JAN-2000.

XX PF 08-JUL-1999; 99WO-NL00436.

XX PR 08-JUL-1998; 98EP-0202297.

XX PA (INTR-) INTROGENE BV.

PI Havenga M, Vogels R, Bout A;

XX WPI; 2000-171149/15.

DR New chimaeric adenoviruses containing a genome derived from different

PT adenovirus serotypes, useful in gene therapy -

PS Example 2; Fig 7; 92pp; English.

XX Y68642-70 represent the amino acid sequences of the fiber proteins

CC of Adenovirus serotypes 8, 9, 13, 14, 20, 23, 24, 25, 27, 28, 29, 30,

CC 32, 33, 34, 35, 36, 37, 38, 39, 42, 43, 44, 45, 46, 47, 48, 49 and 51.

CC The proteins are used in the course of the invention to construct

CC chimaeric adenoviruses with reduced antigenicity. The chimaeric

CC adenoviruses comprise at least part of a fiber protein of an adenovirus

CC serotype providing the chimeric virus with a desired host range and at
 CC least part of a penton or hexon protein from another, less antigenic,
 CC serotype. The chimeric adenoviruses are useful for gene therapy,
 CC especially where repeated delivery is required. Adenoviruses of the
 CC invention are useful can be constructed to have a desired host range and
 CC a diminished capability to raise neutralizing antibodies, an absence of,
 CC or decreased infection of, antigen presenting cells of the immune system
 CC (e.g. macrophages), and an ability to escape trapping in the liver
 CC through increased target cell specificity.
 SQ Sequence 380 AA;

Query Match
 Best Local Similarity 80.5%; Score 33; DB 21; Length 380;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGLLYODG 8
 Db 82 gyltldg 89

RESULT 14

ID Y27892 standard; Protein; 115 AA.
 AC Y27892;

DT 30-JUL-1999 (first entry)
 DE Human secreted protein encoded by gene No. 116.

Human: secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder;
 KW developmental abnormality; foetal deficiency; blood; allergy; leukaemia;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

OS Homo sapiens.
 XX
 PN WO9924836-A1.
 XX
 PD 20-MAY-1999.
 XX
 PF 04-NOV-1998; 98MO-US23435.

PR 17-NOV-1997; 97US-0066100.
 PR 07-NOV-1997; 97US-0064900.
 PR 07-NOV-1997; 97US-0064908.
 PR 07-NOV-1997; 97US-0064911.
 PR 07-NOV-1997; 97US-0064912.
 PR 07-NOV-1997; 97US-0064983.
 PR 07-NOV-1997; 97US-0064984.
 PR 07-NOV-1997; 97US-0064985.
 PR 07-NOV-1997; 97US-0064987.
 PR 17-NOV-1997; 97US-0064988.
 PR 17-NOV-1997; 97US-0066090.
 PR 17-NOV-1997; 97US-0066094.
 PR 17-NOV-1997; 97US-0066095.
 PR 17-NOV-1997; 97US-0066099.

PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Carter KC, Ebner R, Endress GA, Feng P, Janat F;
 PI Kyaw H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;
 PI Ruben SM, Shi Y, Soppet DR, Wei Y;
 XX MPI: 1999-337740/28.
 DR N-PSDB; X85048.
 XX

PT New human secreted proteins and coding sequences useful for treating
 PT disorders of the immune system and hyperproliferative disorders
 PS Disclosure; Page 166; 507pp; English.
 XX

CC This sequence represents a secreted human protein encoded by the gene
 CC clone detailed in the descriptor line. The gene can be used to generate
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc
 CC portion (e.g. X84924) for increasing the stability of the fused protein
 CC as compared to the human protein only.
 CC The invention relates to 125 novel genes and their fragments (nucleic
 CC acid sequences: X84933-X85057; amino acid sequences Y27367-Y27933) which
 CC are useful for preventing, treating or ameliorating medical conditions
 CC e.g. by protein or gene therapy. Also, pathological conditions can be
 CC or by determining the presence of mutations in the new polypeptides in a sample
 CC which tissues they are described for each of the 125 polynucleotides.
 CC (uses).
 CC which tissues they are most highly expressed in (see X84933 for described
 CC uses).

SQ Sequence 115 AA;

Query Match
 Best Local Similarity 78.0%; Score 32; DB 20; Length 115;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGLLYODG 8
 Db 1 gylsvydg 8

RESULT 15

ID Y91324 standard; Protein; 116 AA.
 AC Y91324;

DT 30-MAY-2000 (first entry)
 DE Group B Streptococcus protein sequence SEQ ID NO:53.

Group B Streptococcus; Streptococcus agalactiae; protein antigen;
 KW vaccine; screening; immunogen; detection; diagnosis; infection;
 KW antibody; affibody; antibacterial.
 OS Streptococcus agalactiae.
 XX
 PN WO200006736-A2.
 XX
 PD 10-FEB-2000.
 XX
 PF 27-JUL-1999; 99MO-GB02444.

PR 27-JUL-1998; 98GB-0016335.
 PR 19-MAR-1999; 99US-0125163.
 XX
 PA (MICR-) MICROBIAL TECHNIQUES LTD.
 XX
 PI Le Page RMF, Wells JM, Hanniffy SB;
 PI MPI: 2000-195299/17.
 XX
 PT New Group B Streptococcus protein, useful as vaccine, for diagnosis of
 PT Streptococcal infections and for screening of antibodies or affibodies
 PS Claim 1; Fig 1; 123pp; English.

CC A05803 to A05872 encode proteins, polypeptides and peptides (given in
 CC Y91275 to Y91343) isolated from Group B Streptococcus (GBS), also known
 CC as Streptococcus agalactiae. The GBS polynucleotides and polypeptides
 CC have antibacterial activity. Immunogenic compositions comprising GBS
 CC polynucleotides or polypeptides can be used as vaccines and for the

CC treatment or prophylaxis of GBS infection. The polynucleotides and
CC polypeptides can also be used in the detection of GBS and for screening
CC DNA encoding bacterial cell envelope associated or secreted antigens in
CC gram positive bacteria. A05873 to A05941 represent primers used in the
CC exemplification of the present invention.

XX
SQ Sequence 116 AA;

Query Match 78.0%; Score 32; DB 21; Length 116;
Best Local Similarity 62.5%; Pred. No. 58;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGLVQDC 8
|||::||
Db 90 gglvledg 97

Search completed: June 13, 2001, 14:14:40
Job time: 381 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:10:47 ; Search time 130.61 Seconds
(without alignments)
4.209 Million cell updates/sec

Title: PCT-US01-05825A-20
Perfect score: 41
Sequence: 1 GGLVVDG 8

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	37	90.2	403	2	S42532
2	35	85.4	449	2	F83627
3	35	85.4	508	1	D7EBPH
4	35	85.4	529	1	D7ECBH
5	35	85.4	530	2	G82344
6	35	85.4	532	2	B64100
7	35	85.4	1034	2	T22166
8	34	82.9	427	2	D83347
9	34	82.9	439	2	D70954
10	34	82.9	512	1	D7BSPH
11	34	82.9	1510	2	T33100
12	34	82.9	1542	2	T17459
13	33	80.5	115	1	A46279
14	33	80.5	256	2	A70514
15	33	80.5	319	2	S62191
16	33	80.5	362	2	S37220
17	33	80.5	362	2	S40092
18	33	80.5	606	2	H70816
19	33	80.5	937	2	S7561
20	32	78.0	213	2	A75471
21	32	78.0	305	2	F69947
22	32	78.0	347	1	H64732
23	32	78.0	382	2	H84025
24	32	78.0	420	2	S53916
25	32	78.0	537	2	F70397
26	32	78.0	606	2	A75554
27	32	78.0	757	2	C32052
28	32	78.0	3535	2	E83641
29	32	78.0	5627	2	G83339

30	31	75.6	169	2	E70233	hypothetical prote
31	31	75.6	196	2	F71525	hypothetical prote
32	31	75.6	196	2	H81681	mat protein TC0628
33	31	75.6	198	2	S06176	cytotoxic T-lympho
34	31	75.6	248	1	S01007	granzyme F (EC 3.4
35	31	75.6	248	2	S01006	cytotoxic T-lympho
36	31	75.6	248	2	A33412	cytotoxic T-lympho
37	31	75.6	304	2	B59088	prolyl aminopeptid
38	31	75.6	315	2	S27784	phenylethanolamine
39	31	75.6	394	2	B82000	1-deoxy-D-xylulose
40	31	75.6	394	2	A81229	1-deoxy-D-xylulose
41	31	75.6	396	1	KXB02	plasma protein 2 -
42	31	75.6	516	2	E70779	probable glp1 pro
43	31	75.6	769	2	B83307	probable molybdopt
44	30	73.2	105	2	B64804	YbH protein - Bsc
45	30	73.2	157	2	B70655	2-demethylmenaquin

ALIGNMENTS

RESULT 1
S42532
hypothetical protein 1 - *Synechococcus* sp.
C:Species: *Synechococcus* sp.
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-Mar-1999
C:Accession: S42532
R:Jones, M.C.; Jenkins, J.M.; Smith, A.G.; Howe, C.J.
Plant Mol. Biol. 24, 435-448, 1994
A:Title: Cloning and characterisation of genes for tetrapyrrole biosynthesis from the
A:Reference number: S42531; MUID:94169298
A:Accession: S42532
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-403 <JON>
A>Note: the source is designated as *Anacystis nidulans* R2
C:Superfamily: *Synechococcus* hypothetical protein 1; tetrairicopeptide repeat homology
F:283-316/Domain: tetrairicopeptide repeat homology <TT1>
F:317-350/Domain: tetrairicopeptide repeat homology <TT2>
F:351-384/Domain: tetrairicopeptide repeat homology <TT3>

Query Match
Best Local Similarity 90.2% Score 37; DB 2; Length 403;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGLVVDG 8
DB 215 GGLVVDG 222
RESULT 2
F83627
hypothetical protein PA0142 [imported] - *Pseudomonas aeruginosa* (strain PA01)
C:Species: *Pseudomonas aeruginosa*
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83627
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidis, K.; L
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: F83627
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-449 <STO>
A:Cross-references: GB:AE004452; GB:AE004091; MUD:99945968; PTDN:PA03532.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0142

Query Match 85.4%; Score 35; DB 2; Length 449;
 Best Local Similarity 75.0%; Pred. No. 17;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGLVOD 8
 11111111
 DB 22 GGLVOD 29

RESULT 3
 DTECPH
 purH bifunctional enzyme - Salmonella typhimurium (fragment)
 N:Contains: IMP cyclohydrolase (EC 3.5.4.10); phosphoribosylaminoimidazolecarboxamide fo
 C:Species: Escherichia coli
 C>Date: 30-Sep-1991 #sequence_revision 31-Mar-1993 #text_change 11-Jun-1999
 C:Accession: S18488
 R:Chopra, A.K.; Peterson, J.W.; Prasad, R.
 Biochim. Biophys. Acta 1090: 351-354, 1991
 A:Title: Nucleotide sequence analysis of purH and purD genes from Salmonella typhimurium
 A:Reference number: S18488; MID:92062738
 A:Accession: S18488
 A:Molecule type: DNA
 A:Residues: 1-508 <CHD>
 A:Cross-references: EMBL:M66160; NID:g154286; PIDN:AAA27197.1; PID:g154287
 C:Genetics:
 A:Gene: purH; purJ
 A:Map position: 90 min
 C:Superfamily: purH bifunctional enzyme
 C:Keywords: hydrolase; multifunctional enzyme; purine nucleotide biosynthesis; transfer

Query Match 85.4%; Score 35; DB 1; Length 508;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGLVOD 7
 11111111
 DB 363 GGLVOD 369

RESULT 4
 DTECPH
 purH bifunctional enzyme - Escherichia coli
 N:Contains: IMP cyclohydrolase (EC 3.5.4.10); phosphoribosylaminoimidazolecarboxamide fo
 C:Species: Escherichia coli
 C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 11-Jun-1999
 C:Accession: B34193; S09571; A65208
 R:Alba, A.; Mizobuchi, K.
 J. Biol. Chem. 264: 2123-2126, 1989
 A:Title: Nucleotide sequence analysis of genes purH and purD involved in the de novo pur
 A:Reference number: A92739; MID:90078227
 A:Accession: B34193
 A:Molecule type: DNA
 A:Residues: 1-529 <AIR>
 A:Cross-references: EMBL:J05126; NID:g147419; PIDN:AAA24454.1; PID:g147420
 R:Plamagan, K.A.; Hennigan, S.H.; Vogelbacker, H.H.; Gots, J.S.; Smith, J.M.
 Mol. Microbiol. 4: 381-392, 1990
 A:Title: Purine biosynthesis in Escherichia coli K12: structure and DNA sequence studies
 A:Reference number: S09571; MID:90286915
 A:Accession: S09571
 A:Molecule type: DNA
 A:Residues: 1-529 <FLA>
 A:Cross-references: EMBL:X51950; NID:g42594; PIDN:CAA36212.1; PID:g42595
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 A.: Rose, D.J.; Mau, B.; Sho, Y.
 Science 277: 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MID:97426617
 A:Accession: A65208
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-529 <BLAT>
 A:Cross-references: GB:AE000473; GB:U00096; NID:g2367336; PIDN:AAC76980.1; PID:g17904394

A:Experimental source: strain K-12, substrain MG1695
 C:Genetics:
 A:Gene: purH; purJ
 A:Map position: 90 min
 C:Superfamily: purH bifunctional enzyme
 C:Keywords: hydrolase; multifunctional enzyme; purine nucleotide biosynthesis; transf

Query Match 85.4%; Score 35; DB 1; Length 529;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGLVOD 7
 11111111
 DB 384 GGLVOD 390

RESULT 5
 C82344
 phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase VC0276
 C:Species: Vibrio cholerae
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
 C:Accession: C82344
 R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
 Charlson, D.; Ermolaeva, M.D.; Yamanevan, J.; Baas, S.; Qin, H.; Dragoti, I.; Seller
 1.R.R.; Nekrasova, J.; Venter, J.C.; Fraser, C.M.
 Nucleic Acids Res. 28: 477-483, 2000
 A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MID:20406835
 A:Accession: C82344
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-530 <HIT>
 A:Cross-references: GB:AF004116; GB:AF003852; NID:g9654687; PIDN:AF93451.1; GSPDB:GN
 A:Experimental source: serogroup O1, strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC0276
 A:Map position: 1
 C:Superfamily: purH bifunctional enzyme

Query Match 85.4%; Score 35; DB 2; Length 530;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGLVOD 7
 11111111
 DB 385 GGLVOD 391

RESULT 6
 B64100
 purH bifunctional enzyme - Haemophilus influenzae (strain Rd KW20)
 N:Contains: IMP cyclohydrolase (EC 3.5.4.10); phosphoribosylaminoimidazolecarboxamide
 C:Species: Haemophilus influenzae
 C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 18-Jun-1999
 C:Accession: B64100
 R:Feilschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kervatage
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.; Glodex, A.; Kelley, J.M.; Weidman
 , D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MID:95350630
 A:Accession: B64100
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-532 <TIGR>
 A:Cross-references: GB:U32770; GB:I42023; NID:g1573898; PIDN:AAC22544.1; PID:g1573904
 A:Superfamily: purH bifunctional enzyme
 C:Keywords: hydrolase; multifunctional enzyme; purine nucleotide biosynthesis; transf

Query Match 85.4%; Score 35; DB 2; Length 532;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGLVQD 7
 |||||
 Db 387 GGLVQD 393

RESULT 7
 722166
 hypochelical protein F44D12.1 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T22166
 R:Colles, L.
 submitted to the EMBL Data Library, December 1995
 A:Reference number: Z19525
 A:Accession: T22166
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1034 <MTL>
 A:Cross-references: EMBL:Z68298; PIDN:CAA92607.1; GSPDB:GN00022; CESP:F44D12.1
 A:Experimental source: clone F44D12
 C:Genetics:
 A:Gene: CESP:F44D12.1
 A:Map position: 4
 A:introns: 46/2; 105/3; 187/3; 283/1; 365/3; 444/2; 501/2; 605/1; 688/1; 922/2; 980/3; 1

Query Match 85.4%; Score 35; DB 2; Length 1034;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GGLVQD 8
 |||||
 Db 734 GGLVQD 740

RESULT 8
 D83347
 probable aminotransferase PA2394 [imported] - *Pseudomonas aeruginosa* (strain PA01)
 C:Species: *Pseudomonas aeruginosa*
 C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: D83347
 R:Stover, C.K.; Plam, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
 A:Reference number: A82950; MUID:20437337
 A:Accession: D83347
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-427 <STO>
 A:Cross-references: GB:AE004665; GB:AE004091; NID:99948426; PIDN:AA605782.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA2394

Query Match 82.9%; Score 34; DB 2; Length 427;
 Best Local Similarity 87.5%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGLVQD 8
 |||||
 Db 131 GGLVQD 138

RESULT 9
 D70954
 hypochelical glycine-rich protein RV3595c - *Mycobacterium tuberculosis* (strain H37RV)

C:Species: *Mycobacterium tuberculosis*
 C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C:Accession: D70954
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
 J.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete geno
 A:Reference number: A70500; MUID:98295987
 A:Accession: D70954
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-439 <COL>
 A:Cross-references: GB:Z95557; GB:AL123456; NID:93242276; PIDN:CAB08933.1; PID:921139
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: RV3595c
 C:Superfamily: Phageolus glycine-rich cell wall protein 1.8

Query Match 82.9%; Score 34; DB 2; Length 439;
 Best Local Similarity 87.5%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGLVQD 8
 |||||
 Db 302 GGLVQD 309

RESULT 10
 D7B5PH
 purH bifunctional enzyme - *Bacillus subtilis*
 N:Contains: IMP cyclohydrolase (EC 3.5.4.10); phosphoribosylaminoimidazolecarboxamide
 C:Species: *Bacillus subtilis*
 C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jun-2000
 C:Accession: A29183; F69684
 R:Ebdole, D.J.; Zalkin, H.
 J. Biol. Chem. 262, 8274-8287, 1987
 A:Title: Cloning and characterization of a 12-gene cluster from *Bacillus subtilis* enc
 A:Reference number: A29326; MUID:87250425
 A:Accession: A29183
 A:Molecule type: DNA
 A:Residues: 1-512 <EB>
 A:Cross-references: EMBL:J02732; NID:q143363; PIDN:AA22683.1; PID:q143373
 R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
 C.; Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
 lech, J.; Harwood, C.R.; Henuit, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
 Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
 Y.; M.; Ogawa, K.; Ogilwara, A.; Oudaga, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadle, Y.; Sato, T.; Scanl
 A:Authors: Schell, S.; Schroeder, R.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Se
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpiltra, P.; Tognoni, A.; Tosato, V.; Uchiya
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*
 A:Reference number: A69580; MUID:98044033
 A:Accession: F69684
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-512 <KUN>
 A:Cross-references: GB:Z99107; GB:AL009126; NID:92632866; PIDN:CAB12472.1; PID:926329
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: purH; purH
 A:Map position: 18 min
 C:Superfamily: purH bifunctional enzyme
 C:Keywords: hydrolase; multifunctional enzyme; purine nucleotide biosynthesis; transf

```

Query Match      82.9%; Score 34; DB 1; Length 512;
Best Local Similarity 85.7%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGLVODG 7
DB 370 GGLVODG 376

RESULT 11
hypothetical protein H1E01.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
C:Accession: T33100
R:Giselsel, C.; Mamsley, P.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid H1E01.
A:Reference number: 221282
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1510 <GEN>
A:Cross-references: EMBL:AF067222; PIDN:AAC17017.1; GSPDB:GN00026; CESP:H1E01.3
A:Experimental source: strain Bristol N2; clone H1E01
C:Genetics:
A:Gene: CESP:H1E01.3
A:Map position: X
A:Introns: 73/2; 120/1; 178/3; 227/2; 262/1; 300/2; 345/1; 1430/3; 1468/2

Query Match      82.9%; Score 34; DB 2; Length 1510;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGLVODG 8
DB 1495 GGLVODG 1502

RESULT 12
polyprotein - tomato
C:Species: Lycopersicon esculentum (tomato)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
C:Accession: T17459
R:Parniske, M.; Jones, J.D.
Proc. Natl. Acad. Sci. U.S.A. 96, 5850-5855, 1999
A>Title: Recombination between diverged clusters of the tomato Cf-9 plant disease resist
A:Reference number: 218801; MUID:99254130
A:Accession: T17459
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1542 <PAR>
A:Cross-references: EMBL:AF119040; NID:g4235640; PID:g4235644; PIDN:AA013304.1

Query Match      82.9%; Score 34; DB 2; Length 1542;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGLVODG 8
DB 920 GGLVODG 927

RESULT 13
guanylin precursor [validated] - human
C:Species: Homo sapiens (man)
C>Date: 21-Sep-1993 #sequence_revision 26-May-1995 #text_change 08-Dec-2000
C:Accession: A46279; S29228; S29807

```

```

R:de Sauvage, F.J.; Keshav, S.; Kuang, W.J.; Gillett, N.; Henzel, W.; Goeddel, D.V.
Proc. Natl. Acad. Sci. U.S.A. 89, 9089-9093, 1992
A>Title: Precursor structure, expression, and tissue distribution of human guanylin
A:Reference number: A46279; MUID:93028409
A:Accession: A46279
A:Molecule type: mRNA
A:Residues: 1-115 <DEI>
A:Cross-references: GB:M95174; NID:g306823; PIDN:AAA58625.1; PID:g306824
A:Note: sequence extracted from NCBI backbone (NCBIN:115377, NCBI:P:115376)
R:Wiegand, R.C.; Kato, J.; Huang, M.D.; Fok, K.F.; Kachur, J.F.; Currie, M.G.
FEBS Lett. 311, 150-154, 1992
A>Title: Human guanylin: cDNA isolation, structure, and activity.
A:Reference number: S29228; MUID:93011964
A:Accession: S29228
A:Molecule type: mRNA
A:Residues: 1-115 <WIE>
A:Cross-references: GB:M97496; NID:g183414; PIDN:AAA35915.1; PID:g183415
R:Kuhn, M.; Ralda, M.; Adermann, K.; Schultze-Knappe, P.; Gerzer, R.; Helm, J.M.; Foxes
FEBS Lett. 318, 205-209, 1993
A>Title: The circulating bioactive form of human guanylin is a high molecular weight
A:Reference number: S29807; MUID:93178628
A:Accession: S29807
A:Molecule type: Protein
A:Residues: 22-68 <KUH>
A:Experimental source: Plasma
A:Note: amino-terminal sequencing of mature form and molecular weight of mature form
C:Comment: Guanylin is an endogenous ligand for an intestine-specific receptor guany
C:Genetic: same receptor.
C:Genetic: CDB:GUCR2
A:Cross-references: CDB:136460; OMIM:139392
A:Map position: 1895-1934
C:Superfamily: 1895-1934
C:Keywords: 1895-1934; Intestine
F:12-115/Product: guanylin #status experimental <KAT>

Query Match      80.5%; Score 33; DB 1; Length 115;
Best Local Similarity 75.0%; Pred. No. 10;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGLVODG 8
DB 20 GGLVODG 27

RESULT 14
hypothetical glycine-rich protein Rv2126c - Mycobacterium tuberculosis (strain H37RV)
A:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: A70514
R:Colie, S.T.; Brosch, R.; Parthill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
Rejndean, M.A.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Nature 393, 537-544, 1998
A:Authors: Squires, R.; Stinson, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:9825387
A:Accession: A70514
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-256 <COL>
A:Cross-references: GB:Z97559; GB:AL123456; NID:g3261820; PIDN:CA010721.1; PID:g32618
A:Experimental source: Strain H37RV
C:Genetic:
C:Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homolo

```

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGLLVODG 8

Db 118 GGLLICDG 125

RESULT 15

S62191

5-methyltetrahydropteroyltriglutamate--homocysteine S-methyltransferase (EC 2.1.1.14) -

N:Alternate names: cobalamin-independent methionine synthase

C:Species: Methanobacterium thermoautotrophicum

C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000

C:Accession: S62191

R:Vaupel, M.; Dietz, H.; Linder, D.; Thauer, R.K.

Eur. J. Biochem. 236, 294-300, 1996

A:Title: Primary structure of cyclohydroxylase (Mch) from Methanobacterium thermoautotroph

A:Reference number: S62190; MUID:96184910

A:Accession: S62191

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-319 <VAU>

A:Cross-references: EMBL:X92082; NID:g1103706; PIDN:CAA63062.1; PID:g1103707

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995

C:Genetics:

A:Gene: mchE

A:Start codon: TTG

C:Superfamily: 5-methyltetrahydropteroyltriglutamate--homocysteine S-methyltransferase

C:Keywords: methyltransferase

QY 1 GGLLVODG 8

Db 80 GGMVVKDG 87

Query Match 80.5%; Score 33; DB 2; Length 319;
Best Local Similarity 62.5%; Pred. No. 31;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Search completed: June 13, 2001, 14:10:48

Job time: 149 sec

Page Blank (uspto)

Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:21:48 ; Search time 74.44 Seconds
(without alignments)
3.681 Million cell updates/sec

Title: PCT-US01-05825A-20

Perfect score: 41

Sequence: 1 GGLVQDG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	37	90.2	403 1 YCOA_SYNP7	P42460 synchococc
2	35	85.4	508 1 PUR9_SALTY	P26978 s bifunctio
3	35	85.4	529 1 PUR9_ECOLI	P15639 s bifunctio
4	35	85.4	532 1 PUR9_HAETN	P43852 h bifunctio
5	34	82.9	512 1 PUR9_BACSU	P12048 b bifunctio
6	33	80.5	115 1 GUAN_HUMAN	Q02747 homo sapien
7	33	80.5	309 1 METE_METTM	P55299 methanobact
8	33	80.5	362 1 FIBP_ADE09	P15484 escherichia
9	33	80.5	937 1 CS32_ECOLI	P15484 escherichia
10	32	78.0	305 1 YOKK_BACSU	P45927 bacillus su
11	32	78.0	346 1 GUAQ_ECOLI	P15344 escherichia
12	32	78.0	420 1 SUN4_YEAST	P53616 saccharomyc
13	32	78.0	757 1 FIXI_RHIME	P18398 rhizobium m
14	31	75.6	248 1 GRAE_MOUSE	P08884 mus musculu
15	31	75.6	248 1 GRAF_MOUSE	P08883 mus musculu
16	31	75.6	248 1 GRAG_MOUSE	P13366 mus musculu
17	31	75.6	273 1 YKA2_CABEL	P14254 caenorhabdi
18	31	75.6	360 1 PLSX_DEIRA	Q46578 deinococcus
19	31	75.6	386 1 ICEA_XENLA	P55865 xenopus lae
20	31	75.6	394 1 DXR_NEIMA	Q91x33 neisseria m
21	31	75.6	394 1 DXR_NEIMA	Q91x33 neisseria m
22	31	75.6	396 1 PRIZ_BOVIN	P00744 bos taurus
23	31	75.6	516 1 GLPD_MYCTU	O10502 mycobacteri
24	31	75.6	877 1 SECA_GUTTH	O78441 guillardia
25	30	73.2	182 1 YCBQ_ECOLI	P75855 escherichia
26	30	73.2	247 1 YBFG_ECOLI	P37003 escherichia
27	30	73.2	247 1 GRAB_HUMAN	P10144 h granzyme
28	30	73.2	248 1 GRAD_MOUSE	P11033 mus musculu
29	30	73.2	362 1 MURG_STRCU	O944h4 streptomyc
30	30	73.2	364 1 MURG_STRCU	O942b5 streptomyc
31	30	73.2	370 1 E13B_PEA	O03467 pisum sativ
32	30	73.2	392 1 NOLC_RHIFR	P26508 rhizobium f
33	30	73.2	396 1 A2BP_MOUSE	O91j43 mus musculu

34	30	73.2	396 1 AGS_AGRH	P27875 agrobacteri
35	30	73.2	397 1 A2BP_HUMAN	O9nwb1 homo sapien
36	30	73.2	478 1 NDON_RHOCA	P50973 rhodobacter
37	30	73.2	507 1 ATPA_TOBAC	P00823 nicotiana t
38	30	73.2	543 1 FIBP_ADECC	O65961 canine aden
39	30	73.2	543 1 FIBP_ADECC	P22230 canine aden
40	30	73.2	543 1 FIBP_ADECC	O96689 canine aden
41	30	73.2	590 1 VPP_BPP2	P23479 bacterioph
42	30	73.2	757 1 RRP1_IANN	P21426 influenza a
43	30	73.2	757 1 RRP1_IANN	P16502 influenza a
44	30	73.2	757 1 RRP1_IANN	O82571 influenza a
45	30	73.2	757 1 RRP1_IANN	P16503 influenza a

ALIGNMENTS

RESULT 1	YCOA_SYNP7	STANDARD:	PRT: 403 AA.
ID	YCOA_SYNP7		
AC	P42460:		
DR	01-NOV-1995 (Rel. 32, Created)		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	15-JUL-1998 (Rel. 36, Last annotation update)		
DE	HYPOTHETICAL 45.0 KDA PROTEIN IN COBA 5' REGION.		
OS	Synchococcus sp. (strain PCC 7942) (Anaerobium nidulans R2).		
OC	Bacteria: Cyanobacteria; Chroococcales; Synchococcus.		
OX	NCBI_Taxid=1140;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE-94169298; PubMed-8123787;		
RA	Jones M.C., Jenkins J.M., Smith A.G., Howe C.J.:		
RT	"Cloning and characterization of genes for tetrapyrrole biosynthesis		
RT	from the cyanobacterium Anaerobium nidulans R2."		
RL	Plant Mol. Biol. 24:435-448(1994).		
CC	-1- SIMILARITY: IN THE N-TERMINAL, STRONG TO H. INFLUENZAE HI0653.		
CC	-1- SIMILARITY: CONTAINS 5 TPR REPEATS.		
CC	-1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-14 IS THE INITIATOR.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL: X70966; CAA50301.1; -		
DR	InterPro: IPR001173; -		
DR	InterPro: IPR001440; -		
DR	Pfam: PF00535; Glycos_transf_2; 1.		
DR	Pfam: PF00515; TPR; 1.		
KW	Hypothetical protein; Repeat; TPR repeat.		
FT	REPEAT 208 243		
FT	REPEAT 244 282		
FT	REPEAT 283 316		
FT	REPEAT 317 350		
FT	REPEAT 351 387		
FT	TPR 5.		
SQ	SEQUENCE 403 AA; 44998 MW; 4C0AB39B54988A13 CRC64;		

Query Match Score 37; DB 1; Length 403;

Best Local Similarity 87.58; Pred. No. 3;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGLVQDG 8
|||||:|

DB 215 GGLVQDG 222

RESULT 2
PUR9_SALTY
ID PUR9_SALTY STANDARD: PRT: 508 AA.

```

AC P26978;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE BIFUNCTIONAL PURINE BIOSYNTHESIS PROTEIN PURH [INCLUDES:
DE PHOSPHORIBOSYLIMIDAZOLECARBOXAMIDE FORMYLTRANSFERASE (EC 2.1.2.3)
DE (AICAR TRANSFORMYLASE); IMP CYCLOHYDROLASE (EC 3.5.4.10) (INOSINICASE)
DE (IMP SYNTHETASE) (ATIC)] (FRAGMENT).
GN PURH.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92062738; PubMed=1954258;
RA Chopra A.K., Peterson J.W., Prasad R.;
RT "Nucleotide sequence analysis of purh and purd genes from Salmonella
RT typhimurium."
RL Biochim. Biophys. Acta 1090:351-354(1991).
CC -1- CATALYTIC ACTIVITY: 10-FORMYLTERAHYDROFOLATE + 5'-PHOSPHORIBOSYL-
CC 5'-AMINO-4-IMIDAZOLECARBOXAMIDE = TETRAHYDROFOLATE +
CC 5'-PHOSPHORIBOSYL-5-FORMAMIDO-4-IMIDAZOLECARBOXAMIDE.
CC -1- CATALYTIC ACTIVITY: IMP + H(2)O = 5-FORMAMIDO-1-(5'-PHOSPHORIBOSYL)
CC IMIDAZOLE-4-CARBOXAMIDE.
CC -1- PATHWAY: NINTH AND TENTH STEP IN DE NOVO PURINE BIOSYNTHESIS.
CC -1- DOMAIN: THE IMP CYCLOHYDROLASE ACTIVITY RESIDES IN THE N-TERMINAL
CC REGION (BY SIMILARITY)
CC -1- SIMILARITY: BELONGS TO THE PURH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: M66160; AAA27197.1;
DR PIR: S18488; PREBPH.
DR Strydom: SG10320; PURH.
DR InterPro: IPR002695.
DR Pfam: PF01808; AICARFT__IMCHas; 1.
DR Purine biosynthesis; Transferase; Hydrolase; Multifunctional enzyme.
KW NON_TER
SQ SEQUENCE 508 AA: 55302 MW: 5816A7C2737F152 CRC64;

Query Match 85.4%; Score 35; DB 1; Length 508;
Best Local Similarity 100.0%; Pred. No. 10; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

OY 1 GGLVOD 7
DB 363 GGLVOD 369

RESULT 3
PURH_ECOLI STANDARD: PRT: 529 AA.
AC P15639;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE BIFUNCTIONAL PURINE BIOSYNTHESIS PROTEIN PURH [INCLUDES:
DE PHOSPHORIBOSYLIMIDAZOLECARBOXAMIDE FORMYLTRANSFERASE (EC 2.1.2.3)
DE (AICAR TRANSFORMYLASE); IMP CYCLOHYDROLASE (EC 3.5.4.10) (INOSINICASE)
DE (IMP SYNTHETASE) (ATIC)]
GN PURH.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90078227; PubMed=2687276;
RA Alta A., Mizobuchi K.;
RT "Nucleotide sequence analysis of genes purh and purd involved in the
RT de novo purine nucleotide biosynthesis of Escherichia coli."
RL J. Biol. Chem. 264:21239-21246(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90286915; PubMed=2192220;
RA Flannigan K.A., Hemmigan S.H., Vogelbacker H.H., Gots J.S.;
RT "Purine biosynthesis in Escherichia coli K12: structure and DNA
RT sequence studies of the purd locus."
RL Mol. Microbiol. 4:381-392(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=K12; MG1655;
RX MEDLINE=9408392; PubMed=8265357;
RA Bletcher F.R., Burland V.D., Plunkett G. III, Sofia H.J.;
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RT region from 89.2 to 92.8 minutes."
RL Nucleic Acids Res. 21:5408-5417(1993).
RN [4]
RP SEQUENCE OF 1-12.
RX STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.V., Robinson K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12."
RL Electrophoresis 18:1259-1313(1997).
CC -1- CATALYTIC ACTIVITY: 10-FORMYLTERAHYDROFOLATE + 5'-PHOSPHORIBOSYL-
CC 5'-AMINO-4-IMIDAZOLECARBOXAMIDE = TETRAHYDROFOLATE +
CC 5'-PHOSPHORIBOSYL-5-FORMAMIDO-4-IMIDAZOLECARBOXAMIDE.
CC -1- CATALYTIC ACTIVITY: IMP + H(2)O = 5-FORMAMIDO-1-(5'-PHOSPHORIBOSYL)
CC IMIDAZOLE-4-CARBOXAMIDE.
CC -1- PATHWAY: NINTH AND TENTH STEP IN DE NOVO PURINE BIOSYNTHESIS.
CC -1- DOMAIN: THE IMP CYCLOHYDROLASE ACTIVITY RESIDES IN THE N-TERMINAL
CC REGION (BY SIMILARITY)
CC -1- SIMILARITY: BELONGS TO THE PURH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: J05126; AAA24454.1;
DR EMBL: X51950; CAA36212.1;
DR EMBL: U00006; AAC43104.1;
DR EMBL: AEO00473; AAC76980.1;
DR PIR: B34193; DTECPH.
DR EcoGene: Ecol0795; purh.
DR InterPro: IPR002695;
DR Pfam: PF01808; AICARFT__IMCHas; 1.
DR Purine biosynthesis; Transferase; Hydrolase; Multifunctional enzyme.
SQ SEQUENCE 529 AA: 57329 MW: DC034ED01915DA68 CRC64;

Query Match 85.4%; Score 35; DB 1; Length 529;
Best Local Similarity 100.0%; Pred. No. 10; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

OY 1 GGLVOD 7
DB 384 GGLVOD 390

RESULT 4

```

```

PUR9_HAEIN STANDARD: PRT: 532 AA.
ID PUR9_HAEIN
AC P43832.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE BIFUNCTIONAL PURINE BIOSYNTHESIS PROTEIN PURH [INCLUDES:
DE BIFUNCTIONAL PURINE BIOSYNTHESIS PROTEIN PURH [INCLUDES:
DE PHOSPHORIBOSYLIMIDAZOLECARBOXAMIDE FORMYLTRANSFERASE (EC 2.1.2.3)
DE (ATCAR TRANSFORMYLASE); IMP CYCLOHYDROLASE (EC 3.5.4.10) (INOSINICASE)
DE (IMP SYNTHETASE) (ATC1)].
GN PURH OR H10887.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
CC Haemophilus influenzae.
OK NCBI_Taxid=727;
RX NCBI_Taxid=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kertevage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McEweney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saubol D.M., Brandon R.C.,
RA Fine L.D., Fitchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT Influenzae Rd."
RL Science 269:496-512(1995).
CC -1- CATALYTIC ACTIVITY: 10-FORMYLTERAHYDROFOLATE + 5'-PHOSPHORIBOSYL-
CC 5'-AMINO-4-IMIDAZOLECARBOXAMIDE = TERAHYDROFOLATE +
CC 5'-PHOSPHORIBOSYL-5-FORMAMIDO-4-IMIDAZOLECARBOXAMIDE.
CC -1- CATALYTIC ACTIVITY: IMP + H(2O) = 5-FORMAMIDO-1-(5-PHOSPHORIBOSYL)
CC IMIDAZOLE-4-CARBOXAMIDE.
CC -1- PATHWAY: NINTH AND TENTH STEP IN DE NOVO PURINE BIOSYNTHESIS.
CC -1- DOMAIN: THE IMP CYCLOHYDROLASE ACTIVITY RESIDES IN THE N-TERMINAL
CC REGION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PURH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32770; AAC22544.1; -.
DR TIGR: H10887.
DR InterPro: IPR002695; -.
DR Pfam: PF01808; AICARFT_IMPCHas: 1.
KW Purine biosynthesis; Transferrase; Hydrolase; Multifunctional enzyme.
SQ SEQUENCE 532 AA; 58349 MW; 9DE1E241DD238E87 CRC64;

```

```

Query Match 85.4%; Score 35; DB 1; Length 532;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGLVOD 7
DB 387 GGLVOD 393

```

```

RESULT 5
PUR9_BACSU STANDARD: PRT: 512 AA.
ID PUR9_BACSU
AC P12048.
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)

```

```

DE BIFUNCTIONAL PURINE BIOSYNTHESIS PROTEIN PURH [INCLUDES:
DE PHOSPHORIBOSYLAMINOIMIDAZOLECARBOXAMIDE FORMYLTRANSFERASE (EC 2.1.2.3)
DE (ATCAR TRANSFORMYLASE); IMP CYCLOHYDROLASE (EC 3.5.4.10) (INOSINICASE)
DE (IMP SYNTHETASE) (ATC1)].
GN PURH OR PURH.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
CC Bacillus/Staphylococcus group; Bacillus.
OK NCBI_Taxid=1423;
RX NCBI_Taxid=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=87250425; PubMed=3036807;
RA Eppole D.J., Zalkin H.;
RT "Cloning and characterization of a 12-gene cluster from Bacillus
RT subtilis encoding nine enzymes for de novo purine nucleotide
RT synthesis."
RT J. Biol. Chem. 262:8274-8287(1987).
RN [2]
RP SEQUENCE OF 432-512 FROM N.A.
RC STRAIN=168;
RX MEDLINE=97124186; PubMed=8869499;
RA Boriss R., Portwollik S., Schroeter R.;
RT "The 52 degrees-55 degrees segment of the Bacillus subtilis
RT chromosome: a region devoted to purine uptake and metabolism, and
RT containing the genes cofA, gap and gua and the pur gene cluster
RT within a 34960 bp nucleotide sequence."
RL Microbiology 142:3027-3031(1996).
CC -1- CATALYTIC ACTIVITY: 10-FORMYLTERAHYDROFOLATE + 5'-PHOSPHORIBOSYL-
CC 5'-AMINO-4-IMIDAZOLECARBOXAMIDE = TERAHYDROFOLATE +
CC 5'-PHOSPHORIBOSYL-5-FORMAMIDO-4-IMIDAZOLECARBOXAMIDE.
CC -1- CATALYTIC ACTIVITY: IMP + H(2O) = 5-FORMAMIDO-1-(5-PHOSPHORIBOSYL)
CC IMIDAZOLE-4-CARBOXAMIDE.
CC -1- PATHWAY: NINTH AND TENTH STEP IN DE NOVO PURINE BIOSYNTHESIS.
CC -1- DOMAIN: THE IMP CYCLOHYDROLASE ACTIVITY RESIDES IN THE N-TERMINAL
CC REGION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PURH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J02732; AAA22683.1; -.
DR EMBL: AF011544; AAB72185.1; -.
DR EMBL: Z99107; CAB12472.1; -.
DR PIR: A29183; DTBSPH.
DR Subtilist: BG10710; purh.
DR InterPro: IPR002695; -.
DR Pfam: PF01808; AICARFT_IMPCHas: 1.
KW Purine biosynthesis; Transferrase; Hydrolase; Multifunctional enzyme.
SQ SEQUENCE 512 AA; 55739 MW; 7FD6B0FE548534C CRC64;

```

```

Query Match 82.9%; Score 34; DB 1; Length 512;
Best Local Similarity 85.7%; Pred. NO. 16;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGLVOD 7
DB 370 GGLVOD 376

```

```

RESULT 6
GUAN_HUMAN STANDARD: PRT: 115 AA.
ID GUAN_HUMAN
AC Q02747.
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DT GUANYLIN PRECURSOR (GUANYLATE CYCLASE ACTIVATOR 2A).

```

GN GUCA2A OR GUCA2.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 CC NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Duodenum;
 RX MEDLINE-93011964; PubMed-1327879;
 RA Wiegand R.C., Kato J., Huang M.D., Fok K.F., Kachur J.F.,
 RT Currie M.G.;
 RL "Human guanylin: cDNA isolation, structure, and activity.";
 FBS Lett. 311:150-154(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ileum;
 RX MEDLINE-93028409; PubMed-1409606;
 RA de Sauvage F.J., Keshav S., Kuang W.J., Gillett N., Henzel W.,
 RT Goeddel D.V.;
 RL "Precursor structure, expression, and tissue distribution of human
 guanylin.";
 Proc. Natl. Acad. Sci. U.S.A. 89:9089-9093(1992).
 RN [3]
 RP SEQUENCE OF 22-68.
 RX MEDLINE-93178628; PubMed-8095028;
 RA Khm M., Raide M., Adernann K., Schulz-Knappe P., Gerzer R.,
 RA Heim J.-M., Forsmann W.-G.;
 RT "The circulating bioactive form of human guanylin is a high molecular
 weight peptide (10.3 kDa)";
 FBS Lett. 318:205-209(1993).
 RN [4]
 RP STRUCTURE BY NMR OF 101-115.
 RX MEDLINE-95034794; PubMed-7947768;
 RA Skelton N.J., Garcia K.C., Goeddel D.V., Quan C., Burnier J.P.;
 RT "Determination of the solution structure of the peptide hormone
 guanylin: observation of a novel form of topological
 stereochemistry.";
 Biochemistry 33:13581-13592(1994).
 CC -1- FUNCTION: ENDGENOUS ACTIVATOR OF INTESTINAL GUANYLATE CYCLASE.
 CC IT STIMULATES THIS ENZYME THROUGH THE SAME RECEPTOR BINDING REGION
 CC AS THE HEAT-STABLE ENTEROTOXINS.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN ILEUM AND COLON.
 CC -1- SIMILARITY: BELONGS TO THE GUANYLIN FAMILY. ALSO SIMILAR TO HEAT-
 CC STABLE ENTEROTOXINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M97496; AAA35915.1;
 DR EMBL: M95174; AA58625.1;
 DR PIR: A6279; A6279;
 DR PIR: S29228; S29228;
 DR PDB: 1GNA; 30-SEP-94.
 DR PDB: 1GNB; 30-SEP-94.
 DR MM: 139392; ---
 DR InterPro: IPR000879; ---
 DR Pfam: PF02058; Guanylin; 1.
 DR PRINTS: PRO0774; GUANYLIN.
 KW SIGNAL; 3D-structure.
 FT SIGNAL 1 21
 FT CHAIN 22 115 HMW-GUANYLIN.
 FT PEPTIDE 101 115 GUANYLIN.
 FT DISULFID 104 112
 FT DISULFID 107 115
 FT SEQUENCE 115 AA; 12448 MW; C644C03BAFC26FA CRC64;

Query Match 80.5%; Score 33; DB 1; Length 115;

Best Local Similarity 75.0%; Pred. No. 5.6;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGLYVDG 8
 11:1111
 Db 20 GGVTVQDG 27
 RESULT 7
 MEPE.METTM
 ID MEPE.METTM STANDARD; PRT; 309 AA.
 AC P55299;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PROBABLE METHYLCOBALAMIN: HOMOCYSTEINE METHYLTRANSFERASE (EC 2.1.1.-)
 DE (METHIONINE SYNTHASE).
 GN MEPE.
 OS Methanobacterium thermoautotrophicum (strain Marburg / DSM 2333).
 CC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
 CC Methanothermobacter.
 CC NCBI_TaxID:79929;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-96184910; PubMed-8617278;
 RA Vaupel M., Dietz H., Linde D., Thauer R.K.;
 RT "Primary structure of cyclohydrolase (mch) from Methanobacterium
 thermoautotrophicum (strain Marburg) and functional expression of the
 mch gene in Escherichia coli.";
 Eur. J. Biochem. 236:294-300(1996).
 RN [2]
 RP U. Biochem. 236:294-300(1996).
 CC SEQUENCE OF 3-24. AND CHARACTERIZATION.
 RX MEDLINE-9398345; PubMed-10469143;
 RA Schroeder A., Thauer R.K.;
 RT "Methylcobalamin:homocysteine methyltransferase from Methanobacterium
 thermoautotrophicum. Identification as the mch gene product.";
 Eur. J. Biochem. 263:789-796(1999).
 CC -1- FUNCTION: CATALYZES THE TRANSFER OF A METHYL GROUP FROM
 CC METHYLCOBALAMIN AND METHYLCOBALAMIDE TO HOMOCYSTEINE RESULTING IN
 CC METHIONINE FORMATION.
 CC -1- COFACTOR: ZINC; BINDS ONE MOLE PER SUBUNIT (BY SIMILARITY).
 CC -1- PRIMUM: TERMINAL STEP IN THE DE NOVO BIOSYNTHESIS OF METHIONINE.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SIMILARITY: BELONGS TO THE VITAMIN-B12 INDEPENDENT METHIONINE
 CC SYNTHASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X92082; GAA63062.1; ALT_INIT.
 DR InterPro: IPR002629; ---
 DR Pfam: PF01717; Methionine_synth. 1.
 DR Transferase: Methyltransferase; Methionine biosynthesis; Zinc.
 FT METAL 201 201 ZINC (BY SIMILARITY).
 FT METAL 203 203 ZINC (BY SIMILARITY).
 FT METAL 285 285 ZINC (BY SIMILARITY).
 FT METAL 285 285 ZINC (BY SIMILARITY).
 FT SEQUENCE 309 AA; 33532 MW; 522996A39DF1E31A CRC64;

Query Match 80.5%; Score 33; DB 1; Length 309;
 Best Local Similarity 62.5%; Pred. No. 16;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGLYVDG 8
 11:1111
 Db 70 GGMVVKDG 77


```

RESULT 8
FBP_ADE09 STANDARD: PRT: 362 AA.
AC P36846:
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE FIBER PROTEIN.
GN PIV.
OS Human adenovirus type 9, and Human adenovirus type 15.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=10527, 28276;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TYPE 9 / ISOLATE HICKS, AND TYPE 15 / ISOLATE 5399 & MORRISON;
RX MEDLINE=9513193; PubMed=7831811;
RA Piling-Akerblom P., Adriaen T.;
RT "Characterization of adenovirus subgenus D fiber genes.";
RL Virology 206:564-571(1995).
CC -1- FUNCTION: RECOGNIZES THE CELL RECEPTOR, SERVES AS THE LIGAND
CC BETWEEN THE ADENOVIRUS CAPSID AND THE HOST CELL RECEPTOR.
CC -1- SUBUNIT: HOMOTRIMER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X74659; CA52723.1; -
DR EMBL: X74658; CA52722.1; -
DR EMBL: X76706; CA54127.1; -
DR PIR: S37220; S37220.
DR PIR: S37214; S37214.
DR PIR: S40092; S40092.
DR HSP: P1818; 1KMB.
DR InterPro: IPR000331; -
DR InterPro: IPR000339; -
DR InterPro: IPR000978; -
DR Pfam: PF00608; adeno_fiber2; 2.
DR Pfam: PF00541; adeno_fiber; 1.
DR PRINTS: PR00307; ADENOVSFIBRE.
KM Fiber protein.
SQ SEQUENCE 362 AA; 39420 MW; 2662E5546A9C4DA0 CRC64;

Query Match 80.5%; Score 33; DB 1; Length 362;
Best Local Similarity 75.0%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGLVVDG 8
DB 67 GGLTLDG 74

RESULT 9
CS32_ECOLI STANDARD: PRT: 937 AA.
AC P15484; P15485; P15487;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE OUTER MEMBRANE USHER PROTEIN CS3-2 PRECURSOR (CS3 PILI SYNTHESIS 104
DE KDA PROTEIN).
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PB176;

```

```

RX MEDLINE=90158116; PubMed=2576094;
RA Jalajakumari M.B., Thomas C.J., Halter R., Manning P.A.;
RT "Genes for biosynthesis and assembly of CS3 pili of CFA/II
RT enterotoxigenic Escherichia coli: novel regulation of pilus
RT production by bypassing an amber codon.";
RL Mol. Microbiol. 3:1685-1695(1989).
CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR THE BIOGENESIS OF
CC MATURE CS3 PILI.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE
CC (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: FIVE PROTEIN ARE PRODUCED BY ALTERNATIVE
CC INITIATION. THE FIFTH PROTEIN REQUIRES THE SUPPRESSION OR
CC READTHROUGH OF AN INTERNAL AMBER CODON IN POSITION 754.
CC -1- SIMILARITY: BELONGS TO THE FIMBRIAL EXPORT USHER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X16944; CA34817.1; -
DR EMBL: X16944; CA34816.1; ALT_INIT.
DR InterPro: IPR000015; -
DR Pfam: PF00577; usher; 1.
DR PROSITE: PS01151; FIMBRIAL_USHER; 1.
KM Outer membrane; Transmembrane; Fimbria; Transport; Signal;
KW Alternative initiation.
FT SIGNAL 1 ?
FT CHAIN 1 ? POTENTIAL.
FT CHAIN 181 937 CS3 PILI SYNTHESIS 104 KDA PROTEIN.
FT CHAIN 181 753 CS3 PILI SYNTHESIS 63 KDA PROTEIN.
FT CHAIN 317 753 CS3 PILI SYNTHESIS 48 KDA PROTEIN.
FT CHAIN 451 753 CS3 PILI SYNTHESIS 33 KDA PROTEIN.
FT CHAIN 572 753 CS3 PILI SYNTHESIS 20 KDA PROTEIN.
SQ SEQUENCE 937 AA; 104150 MW; 6C53056EE96277B5 CRC64;

Query Match 80.5%; Score 33; DB 1; Length 937;
Best Local Similarity 75.0%; Pred. No. 49;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGLVVDG 8
DB 553 GGLVYDG 560

RESULT 10
YOBK_BACSU STANDARD: PRT: 305 AA.
AC P45927;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHETICAL 32.7 KDA PROTEIN IN SPOILITIC-OWLA INTERGENIC REGION.
GN YOBK.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX Takemaru K.-I., Mizuno M., Sato T., Takeuchi M., Kobayashi Y.;
RT "Complete nucleotide sequence of a skin element excised by DNA
RT rearrangement during sporulation in Bacillus subtilis.";
RL Microbiology 141:323-327(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RA Kobayashi Y., Mizuno M., Masuda S., Takemaru K., Hosono S.,

```

```

RA Sato T., Takeuchi M.:
RL Submitted (Mar-1996) to the EMBL/Genbank/DBJ databases.
RM [3]
RP IDENTIFICATION:
RX MEDLINE-96084975; PubMed-7489895;
RT Medigue C., Moszer I., Viari A., Danchin A.;
"Analysis of a Bacillus subtilis genome fragment using a co-operative
computer system prototype."
RL Gene 165:GC37-GC51(1995).
CC -1- SIMILARITY: STRONG, TO B. SUBTILIS XKDK.
CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 179
CC ONWARD AND IS SHORTER (226 AA) DUE TO A FRAMESHIFT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D32216; BAA06943.1; ALT_FRAME.
DR EMBL: D84432; BAA12406.1; -.
DR EMBL: Z99117; CAB14349.1; -.
DR SUBTILIS: BG11282; YQDK.
KW Hypothetical protein.
SQ SEQUENCE 305 AA; 32667 MW; FE12A8987867B1EF CRC64;

Query Match 78.0%; Score 32; DB 1; Length 305;
Best Local Similarity 62.5%; Pred. No. 25;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGLVQDG 8
DB 271 GGVLEEDG 278

RESULT 11
ID GUAC_ECOLI STANDARD; PRT; 346 AA.
AC P15344; P78048;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GMP REDUCTASE (EC 1.6.6.8) (GUANOSINE 5'-MONOPHOSPHATE
DE OXIDOREDUCTASE) (GUANOSINE MONOPHOSPHATE REDUCTASE).
GN GUAC.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12:
RX MEDLINE-89061679; PubMed-2904262;
RT Andrews S.C., Guest J.R.;
"Nucleotide sequence of the gene encoding the GMP reductase of
RT Escherichia coli K12."
RL Biochem. J. 235:35-43(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / W3110;
RX MEDLINE-94261430; PubMed-8203364;
RT Fujita N., Mori H., Yura T., Ishihama A.;
"Systematic sequencing of the Escherichia coli genome: analysis of
RT the 2.4-4.1 min (110,917-193,643 bp) region."
RL Nucleic Acids Res. 22:1637-1639(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RT Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

```

```

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
"the complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE OF 243-346 FROM N.A.
RC STRAIN-K12;
RX MEDLINE-95047556; PubMed-7959070;
RT Whitechurch C.B., Matlick J.S.;
"Escherichia coli contains a set of genes homologous to those
RT involved in protein secretion, DNA uptake and the assembly of type-4
RT fimbriae in other bacteria."
RL Gene 150:9-15(1994).
RN [5]
RP SEQUENCE OF 1-182;
RC STRAIN-K12 / EMG2;
RX MEDLINE-97443975; PubMed-9298646;
RT Link A.J., Robison K., Church G.M.;
"Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12."
RL Electrophoresis 18:1259-1313(1997).
CC -1- FUNCTION: IT FUNCTIONS IN THE CONVERSION OF NUCLEOBASE, NUCLEOSIDE
CC AND NUCLEOTIDE DERIVATIVES OF G TO A NUCLEOTIDES, AND IN
CC MAINTAINING THE INTRACELLULAR BALANCE OF A AND G NUCLEOTIDES.
CC -1- CATALYTIC ACTIVITY: NADPH + GUANOSINE 5'-PHOSPHATE = NADP(+) +
CC INOSINE 5'-PHOSPHATE + NH(3).
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC GMP REDUCTASE
CC AND TO IMP DEHYDROGENASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X07917; CAA30751.1; -.
DR EMBL: D26562; CAB20312.1; -.
DR EMBL: AE000119; AAC73215.1; -.
DR EMBL: L28105; AAC36926.1; -.
DR PIR: S01671; S01671.
DR PIR: S45182; S45182.
DR HSSP: P50097; IAK5.
DR ECO2DBASE: H037.4; 6TH EDITION.
DR EcoGene: EG10422; guac.
DR InterPro: IPR001093; -.
DR Pfam: PF00478; IMPDH_C; 1.
DR Pfam: PF01574; IMPDH_N; 1.
DR PROSITE: PS00487; IMP_DH_GMP_RED; 1.
KW Oxidoreductase; NADP.
FT BINDING 186 186 IMP (POTENTIAL).
FT NP_BIND 216 238 NADP (POTENTIAL).
FT CONFLICT 233 234 AR -> GGG (in Ref. 3).
SQ SEQUENCE 346 AA; 37440 MW; BA890DA702DA3A7 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 346;
Best Local Similarity 62.5%; Pred. No. 28;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGLVQDG 8
DB 213 GGMVSDG 220

RESULT 12
ID SUN4_YEAST STANDARD; PRT; 420 AA.
AC P53616;
DT 01-OCT-1996 (Rel. 34, Created)

```

```

DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel.37, Last annotation update)
DE PROTEASOME COMPONENT SUN4
CN SUN4 OR YNL066W OR N2411 OR YNL211W
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
ON NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / F1676;
RX MEDLINE=96021608; PubMed=8533472;
RA Berger P., Dolignon F., Crouzet M.;
RT "The sequence of a 44 420 bp fragment located on the left arm of
RL chromosome XIV from Saccharomyces cerevisiae."
RN [2]
RP ERRATUM.
RX MEDLINE=97060022; PubMed=8904343;
RA Berger P., Dolignon F., Crouzet M.;
RT Yeast 12:297-297(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / FY1679;
RX MEDLINE=96267764; PubMed=8701611;
RA Pochlmann R., Philippson P.;
RT "Sequencing a cosmid clone of Saccharomyces cerevisiae chromosome XIV
RL reveals 12 new open reading frames (ORFs) and an ancient duplication
RN of six ORFs."
RT Yeast 12:391-402(1996).
CC -1- SIMILARITY: BELONGS TO THE NCA3/SUN4/STML/YKRO42W FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U12141; AAA9645.1; -
DR EMBL; X86470; CAA60196.1; -
DR EMBL; Z71342; CAA95939.1; -
DR SGD; S0005010; SUN4.
KW Proteasome.
FT DOMAIN 336 POLY-SER.
SO SEQUENCE 420 AA; 43442 MM; F1FB6CD46F2CDA13 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 420;
Best Local Similarity 75.0%; Pred. No. 34;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGLLVQDG 8
   ||||:||
Db 235 GGLLCKDG 242

RESULT 13
FIX1_RHIME STANDARD; PRT; 757 AA.
AC P18398;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NITROGEN FIXATION PROTEIN FIX1 (EL-E2 TYPE CATION ATPASE FIX1)
DE (EC 3.6.1.-).
GN FIX1.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OS Plasmid pSyma (megaplasmid 1).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
ON NCBI_TaxID=382;

```

```

RN [1]
RP [1] SEQUENCE FROM N.A.
RC STRAIN-RCR2011 / SU47;
RX MEDLINE=89123173; PubMed=2536685;
RA Kahn D., David M., Domergue O., Daveran M.-L., Ghai J., Hirsch P.R.,
RA Batut J.;
RA "Rhizobium meliloti fixGHI sequence predicts involvement of a
RT specific cation pump in symbiotic nitrogen fixation.";
RL J. Bacteriol. 171:929-939(1989).
CC -1- FUNCTION: FIXI IS A PUMP OF A SPECIFIC CATION INVOLVED IN
CC SYMBIOTIC NITROGEN FIXATION. THE FOUR PROTEINS FIXG, FIXH, FIXI,
CC AND FIXJ MAY PARTICIPATE IN A MEMBRANE-BOUND COMPLEX COUPLING
CC THE FIXI CATION PUMP WITH A REDOX PROCESS CATALYZED BY FIXG.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + ORTHOPHOSPHATE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
CC (EI-E2 ATPASES). SUBFAMILY IB.
CC -1- SIMILARITY: CONTAINS 1 HEAVY-METAL-ASSOCIATED (HMA) DOMAIN.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC CC at the European Bioinformatics Institute. There are no restrictions on
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: M24144; AAA26273.1; -.
DR EMBL: Z21854; CAI79907.1; -.
DR PIR: C32052; C32052.
DR PIR: S39994; S39994.
DR PIR: S32847; S32847.
DR HSSP: Q27546; IMAS.
DR InterPro: IPR001577; -.
DR InterPro: IPR001934; -.
DR Pfam: PF00122; EI-E2_Atpase; 1.
DR Pfam: PF00403; HMA; 1.
DR PROSITE: PS00154; ATPASE_EI_E2; 1.
DR PROSITE: PS01047; HMA; 1.
KW Hydrolyase; Transmembrane; Phosphorylation; ATP-binding; Metal-binding;
KW Nitrogen fixation; Plasmid.
FT DOMAIN 1 123 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 124 144 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 145 154 POTENTIAL.
FT TRANSMEM 155 175 POTENTIAL.
FT DOMAIN 176 209 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 210 230 POTENTIAL.
FT DOMAIN 231 374 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 375 395 POTENTIAL.
FT DOMAIN 396 397 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 398 418 POTENTIAL.
FT DOMAIN 419 590 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 591 611 POTENTIAL.
FT DOMAIN 612 689 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 690 710 POTENTIAL.
FT DOMAIN 711 711 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 712 732 POTENTIAL.
FT DOMAIN 733 757 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 43 72 HMA.
FT METAL 48 48 POTENTIAL.
FT METAL 51 51 POTENTIAL.
FT MOD_RES 454 454 PHOSPHORYLATION (PROBABLE).
SQ SEQUENCE 757 AA; 79559 MW; DE2E5C6249254AAS CRC64;

Query Match 78.0%; Score 32; DB 1; Length 757;
Best Local Similarity 62.5%; Pred. No. 63;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
DB 432 GGVWVKDG 439

```

```

RESULT 14
ID GRAF_MOUSE STANDARD; PRT; 248 AA.
AC P08884; P97389.
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE GRANZYME E PRECURSOR (EC 3.4.21.-) (CYTOTOXIC CELL PROTEASE 3) (CCP3)
DE (CTL SERINE PROTEASE 2) (D12) (CYTOTOXIC SERINE PROTEASE 2) (MCP2)
DE GZME OR CTIA6 OR CTIA-6 OR CCP3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID:10090.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-68271600; PubMed-3292281;
RA Bleackley R.C., Duggan B., Enman N., Lobe C.G.;
RT "Isolation of two cDNA sequences which encode cytotoxic cell
RT proteases." FEBS Lett. 234:153-159(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX Prendergast J.A., Pinkoski M., Wolfenden A., Bleackley R.C.;
RA Submitted (NOV-1990) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN-129/SV;
RA Pham C.T.N., MacIvor D.M., Hug B.A., Heusel J.W., Ley T.J.;
RN Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RP [4]
RX MEDLINE-8-248 FROM N.A.
RA MEDLINE-88263037; PubMed-3260382;
RX Jenne D.E., Ray C., Haefliger J.-A., Qiao B.-Y., Groscurch P.,
RA Tschopp J.;
RT "Identification and sequencing of cDNA clones encoding the granule-
RT associated serine proteases granzymes D, E, and F of cytolytic T
RT lymphocytes." Proc. Natl. Acad. Sci. U.S.A. 85:4814-4818(1988).
RN [5]
RP SEQUENCE OF 14-248 FROM N.A.
RX TISSUE-Cytotoxic T-cell;
RA MEDLINE-89036001; PubMed-3053963;
RX Kwon B.S., Kestler D., Lee E., Hakulchik M., Young J.D.;
RT "Isolation and sequence analysis of serine protease cDNAs from mouse
RT cytolytic T lymphocytes." J. Exp. Med. 168:1839-1854(1988).
RN [6]
RP SEQUENCE OF 21-40.
RX MEDLINE-87213932; PubMed-355842;
RA Masson D., Tschopp J.;
RT "A family of serine esterases in lytic granules of cytolytic T
RT lymphocytes." Cell 49:679-685(1987).
RN [7]
RP LYSIS IN CELL-MEDIATED IMMUNE RESPONSES.
DE (CTL SERINE PROTEASE 3) (C134) (CYTOTOXIC SERINE PROTEASE 3) (MCP3)
DE GZME OR CTIA7 OR CTIA-7 OR CCP4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID:10090.
RN [11]
RP SEQUENCE FROM N.A.
RX TISSUE-Spleen;
RA MEDLINE-91318146; PubMed-1861068;
RX Jenne D.E., Zimmer M., Garcia-Sanz J.A., Tschopp J.F., Licher P.;
RT "Genomic organization and subcellular localization of the
RT murine granzyme F, a serine protease expressed in CD8+ T cells." J
RT Immunol. 147:1045-1052(1991).
RN [12]
RP SEQUENCE FROM N.A.
RX MEDLINE-91350184; PubMed-1880801;
RA Prendergast J.A., Pinkoski M., Wolfenden A., Bleackley R.C.;
RT "Structure and evolution of the cytotoxic cell proteinase genes CCP3,

```

```

DR EMBL: U66474; AAB19192.1; -
DR EMBL: X56988; A440306.1; -
DR EMBL: J03256; A437737.1; -
DR EMBL: X14093; CAA32255.1; -
DR PIR: S01006; S01006
DR PIR: B26944; B26944
DR PIR: B36172; B36172
DR PIR: S24935; S24935
DR HSP: P04187; SCPL1
DR MEROPS: S01395; SCPL1
DR MG: M61109265; GZME
DR InterPro: IPR001254; -
DR InterPro: IPR001314; -
DR Pfam: PF00089; Trypsin; 1
DR PRINTS: PR00722; CHYMOTRYPSIN
DR PROSITE: PS00134; TRYPsin_HIS_1
DR PROSITE: PS00135; TRYPsin_SER_1
KW Hydrolase; Serine protease; Zymogen; Signal; T-cell; Cytolysis.
FT SIGNAL 1 18
FT PROPEP 19 20
FT CHAIN 21 248
FT ACT_SITE 65 65 GRANZYME E.
FT ACT_SITE 109 109 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 204 204 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFD 50 66 BY SIMILARITY.
FT DISULFD 143 210 BY SIMILARITY.
FT DISULFD 175 189 BY SIMILARITY.
FT CARBOHYD 68 68 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 102 102 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 223 223 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CONFLICT 5 5 L -> P (IN REF. 3).
FT CONFLICT 132 132 R -> K (IN REF. 5).
FT CONFLICT 150 150 S -> P (IN REF. 3 AND 4).
SQ SEQUENCE 248 AA; 27494 MW; 3A31912A5E93D3F CRC64;

```

```

Query Match
Best Local Similarity: 85.78; Pred. No. 32; Length 248;
Matches: 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGLVOD 7
DB 51 GGLVOD 57

```

```

RESULT 15
ID GRAF_MOUSE STANDARD; PRT; 248 AA.
AC P08883;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE GRANZYME F PRECURSOR (EC 3.4.21.-) (CYTOTOXIC CELL PROTEASE 4) (CCP4)
DE (CTL SERINE PROTEASE 3) (C134) (CYTOTOXIC SERINE PROTEASE 3) (MCP3)
DE GZMF OR CTIA7 OR CTIA-7 OR CCP4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID:10090.
RN [11]
RP SEQUENCE FROM N.A.
RX TISSUE-Spleen;
RA MEDLINE-91318146; PubMed-1861068;
RX Jenne D.E., Zimmer M., Garcia-Sanz J.A., Tschopp J.F., Licher P.;
RT "Genomic organization and subcellular localization of the
RT murine granzyme F, a serine protease expressed in CD8+ T cells." J
RT Immunol. 147:1045-1052(1991).
RN [12]
RP SEQUENCE FROM N.A.
RX MEDLINE-91350184; PubMed-1880801;
RA Prendergast J.A., Pinkoski M., Wolfenden A., Bleackley R.C.;
RT "Structure and evolution of the cytotoxic cell proteinase genes CCP3,

```

RT CCP4 and CCP5.";
 RL J. Mol. Biol. 220:867-875(1991).
 RN [3]
 RX SEQUENCE FROM N.A.
 RA MEDLINE=88271600; PubMed=3292281;
 RT Blackley R.C., Duggan B., Ehrman N., Lobe C.G.;
 RT "Isolation of two cDNA sequences which encode cytotoxic cell
 RT proteases.";
 RL FEBS Lett. 234:153-159(1988).
 RN [4]
 RX SEQUENCE FROM N.A.
 RA MEDLINE=88263037; PubMed=3260382;
 RT Jenner D.E., Rey C., Haefliger J.-A., Qiao B.-Y., Groscurch P.,
 RA Tschoop J.;
 RT "Identification and sequencing of cDNA clones encoding the granule-
 RT associated serine proteases granzymes D, E, and F of cytolytic T
 RT lymphocytes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:4814-4818(1988).
 RN [5]
 RX SEQUENCE FROM N.A.
 RA TISSUE=Cytotoxic T-cell;
 RL MEDLINE=89036001; PubMed=3053963;
 RA Kwon B.S., Keastler D., Lee E., Makulchik M., Young J.D.;
 RT "Isolation and sequence analysis of serine protease cDNAs from mouse
 RT cytolytic T lymphocytes.";
 RL J. Exp. Med. 168:1839-1854(1988).
 RN [6]
 RX SEQUENCE OF 21-40.
 RA MEDLINE=87215932; PubMed=3555842;
 RT Maason D., Tschoop J.;
 RT "A family of serine esterases in lytic granules of cytolytic T
 RT lymphocytes.";
 RL Cell 48:679-685(1987).
 RN [7]
 RX SEQUENCE OF 21-45.
 RA MEDLINE=93044519; PubMed=2152187;
 RT Jiang S., Hasseikus-Light C.S., Ojcius D.M., Young J.D.E.;
 RT "Purification of a membrane-associated serine esterase from murine
 RT cytotoxic T lymphocytes by a single reverse-phase column.";
 RL Protein Expr. Purif. 1:77-80(1990).
 CC -1- FUNCTION: THIS ENZYME IS PROBABLY NECESSARY FOR TARGET CELL
 CC LYSIS IN CELL-MEDIATED IMMUNE RESPONSES.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC GRANULES OF CYTOLYTIC
 CC T-LYMPHOCYTES.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY. STRONGEST TO OTHER GRANZYMES AND TO MAST CELL
 CC PROTEASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M36902; AAA37488.1; -
 DR EMBL: X56989; CAA40307.1; -
 DR EMBL: M96930; AAA37741.1; -
 DR EMBL: J03257; AAA37738.1; -
 DR EMBL: X12823; CAA31310.1; -
 DR EMBL: X14094; CAA32256.1; -
 DR PIR: S01007; S01007.
 DR HSSP: P04187; ZCP1.
 DR MEROPS: S01.401; -
 DR MGD: MGI:109254; Gzmf.
 DR InterPro: IPR001254; -
 DR InterPro: IPR001314; -
 DR Pfam: PF00089; Trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; zymogen; signal; T-cell; Cytolysis.

FT	SIGNAL	1	18	
FT	PROPEP	19	20	
FT	CHAIN	21	248	
FT	ACT_SITE	65	65	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	109	109	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	204	204	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	DISULFID	50	66	BY SIMILARITY.
FT	DISULFID	143	210	BY SIMILARITY.
FT	DISULFID	175	189	BY SIMILARITY.
FT	CARBOHYD	106	106	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	154	154	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	223	223	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	248 AA;	27642 MW;	02B4B67F10D0C38 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 248;
 Best Local Similarity 85.7%; Pred. No. 32;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGLVOD 7
 II IIII
 DB 51 GGLVOD 57

Search completed: June 13, 2001, 14:21:48
 Job time: 808 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:20:37 ; Search time 225.85 Seconds
(without alignments)
4.152 Million cell updates/sec

Title: PCT-US01-05825A-20

Perfect score: 41

Sequence: 1 GGLLYQDG 8

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_15:*

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhc:*
- 8: sp.organelle:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp.todent:*
- 12: sp.unclassified:*
- 13: sp.vertebrate:*
- 14: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	85.4	529	2 09L9H7	09L9H7 salmonella
2	35	85.4	530	2 09KVB0	09KVB0 vibrio chol
3	35	85.4	620	10 09SKR3	09SKR3 arabisidopsis
4	35	85.4	1034	5 020398	020398 caenorhabdi
5	35	85.4	1491	10 09XE26	09XE26 oryza sativ
6	34	82.9	439	2 006287	006287 mycobacteri
7	34	82.9	1240	10 09LH01	09LH01 arabidopsis
8	34	82.9	1510	5 061802	061802 caenorhabdi
9	34	82.9	1542	10 09ZS84	09ZS84 lycopersico
10	33	80.5	303	2 P97157	P97157 escherichia
11	33	80.5	365	14 064823	064823 human adeno
12	33	80.5	476	10 004927	004927 human adeno
13	33	80.5	476	10 004927	004927 allium tube
14	33	80.5	606	2 053884	053884 mycobacteri
15	32	78.0	213	2 09RW35	09RW35 mycobacteri
16	32	78.0	382	2 09K8J8	09K8J8 bacillus ha
17	32	78.0	537	2 005459	005459 mycobacteri
18	32	78.0	569	1 074030	074030 mechanobact
19	32	78.0	606	2 09RX21	09RX21 deinococcus

20	32	78.0	708	10 09LMP3	09LMP3 oryza sativ
21	32	78.0	1335	2 09LA58	09LA58 escherichia
22	32	78.0	1335	2 09LA54	09LA54 escherichia
23	31	75.6	40	2 09L171	09L171 streptomyces
24	31	75.6	156	2 087722	087722 caulobacter
25	31	75.6	169	2 050741	050741 borrelia bu
26	31	75.6	196	2 084353	084353 chlamydia t
27	31	75.6	196	2 09PK45	09PK45 chlamydia m
28	31	75.6	241	11 063637	063637 raltus norv
29	31	75.6	242	2 052558	052558 amycolatops
30	31	75.6	250	5 09NN02	09NN02 leishmania
31	31	75.6	304	2 048556	048556 lactobacill
32	31	75.6	304	2 048570	048570 lactobacill
33	31	75.6	394	2 09K1G8	09K1G8 neisseria m
34	31	75.6	421	2 09UX33	09UX33 neisseria m
35	31	75.6	489	10 09L116	09L116 streptomyces
36	31	75.6	528	2 068180	068180 uncultured
37	31	75.6	528	2 068180	068180 uncultured
38	31	75.6	1241	10 09LHK4	09LHK4 oryza sativ
39	31	75.6	1587	10 09SL00	09SL00 oryza sativ
40	31	75.6	2207	5 09U0V2	09U0V2 leishmania
41	30	73.2	84	2 09XBV7	09XBV7 mycobacteri
42	30	73.2	96	2 09R1I6	09R1I6 streptococc
43	30	73.2	119	8 09MUN9	09MUN9 mesosigma
44	30	73.2	152	6 029224	029224 sus scrofa
45	30	73.2	157	2 P96224	P96224 mycobacteri

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	529 AA.
09L9H7	09L9H7	09L9H7:		
AC	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)			
DE	PHOSPHORIBOSYLTRANSFERASE FORMYLTRANSFERASE AND IMP			
DE	CYCLOHYDROLASE.			
GN	PURH.			
OS	Salmonella typhimurium LT2.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Salmonella.			
OX	NCBI_TaxID=99287;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-SGSC1412;			
RA	latrelle, P;			
RT	"The sequence of Salmonella typhimurium fragment STM1 (I-Ceu-F).";			
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-SGSC1412;			
RA	Washu;			
RT	"The Salmonella typhimurium Genome Sequencing Project.";			
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-SGSC1412;			
RA	Waterston R.;			
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF10176; AAF3520.1; -			
SO	SEQUENCE 529 AA; 57467 MW; 5D713504AD644778 CRC64;			

Query Match Score 85.4%; DB 2; Length 529;

Best Local Similarity 100.0%; Pred. No. 58;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGLLYQD 7
DB 384 GGLLYQD 390

```

RESULT 2
OQKV80 PRELIMINARY: PRT: 530 AA.
AC OQKV80:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DE PHOSPHORIBOSYLTAMINOIMIDAZOLICARBOXAMIDE FORMYLTRANSFERASE/IMP
DE CYCLICHYDROLASE.
GN VCO276
OS Bacteria: Proteobacteria: gamma subdivision: Vibrionaceae: Vibrio.
OC NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gatt M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umamya L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Olin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
DR EMBL: AE004115; AAF93451.1;
DR TIGR: VCO276;
KW Transferase.
SQ SEQUENCE 530 AA: 57326 MW: B564A6D4C42FA617 CNC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGLVOD 7
Db 385 GGLVOD 391

RESULT 3
OQSR3 PRELIMINARY: PRT: 620 AA.
AC OQSR3:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
DE 01-MAY-2000 (TREMBlrel. 13, last annotation update)
DE F2103.24 PROTEIN.
GN F2103.24.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Beilto M., Creasy T.H., Haas B.,
RA Rensing C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome IIT BAC F2103 genomic sequence."
RT Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC009853; AAF02157.1;
DR NCBI_TaxID=3702;
SQ SEQUENCE 620 AA: 68199 MW: 1F2D06A4A3972D9E CRC64;

Query Match
Best Local Similarity 85.4%; Score 35; DB 10; Length 620;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 GGLVODG 8
Db 513 GGLVODG 520

RESULT 4
OQ20398 PRELIMINARY: PRT: 1034 AA.
AC OQ20398:
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DE 01-JUN-2000 (TREMBlrel. 14, last annotation update)
DE F44D12.1 PROTEIN.
GN F44D12.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Rhabditinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX SOLES L.;
RX Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RX MEDLINE=94150718; PubMed=7906398;
RA Willford R., Almscough R., Anderson K., Baynes C., Berks M., Coulson A.,
RA Baillif J., Sutton J., Comeff M., Copestey T., Cooper J., Fulton L.,
RA Craxton M., Dutton J., Durrin R., Favello J., Fulton L.,
RA Gardner A., Gnan S., Du Z., Durrin R., Favello J., Fulton L.,
RA Jones M., Jersha J., Kewen J., Kisten J., Lister N., Lattelle F.,
RA Lightning J., Lloyd C., McIntyre A., Mortimore B., O'Callaghan M.,
RA Paisson J., Peltier A., Rinken L., Koopra A., Saunders D., Showman R.,
RA Smalton N., Smith A., Sonnenmeier E., Staeden R., Sulston J.,
RA Thiermy-Mieg Y., Thomas K., Winkler L., Winkler S., Vaughan R., Winkler R.,
RA Watson A., Wellstock L., Winkler S., Winkler S., Winkler P., III of C.
RT elegans "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
DR EMBL: Z68298; CA92607.1;
DR INTERPRO: IPR000619;
DR INTERPRO: IPR001478;
DR PIRAM: PR00595; PDZ: 1;
DR PROSITE: PS50052; GUANYLATE_KINASE_2; 1;
SQ SEQUENCE 1034 AA: 116053 MW: 7E29646264FCD5BC CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 13e+02; Length 1034;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GGLVODG 8
Db 734 GGLVODG 740

RESULT 5
OQX26 PRELIMINARY: PRT: 1491 AA.
AC OQX26:
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, last sequence update)
DE 01-JUN-2000 (TREMBlrel. 14, last annotation update)
DE HYPOHENTICAL 170.8 KDA PROTEIN.
GN Oryza sativa (Rice).
OS Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, PAC

```



```

RT clone:PO026F07.":
RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AP000364; BAA81778.1; -.
DR INTERPRO: IPR000477; -.
DR INTERPRO: IPR001878; -.
DR INTERPRO: IPR001969; -.
DR PFAM: PF000078; rvt; 1.
DR PFAM: PF000098; zfc-CCHC; 1.
DR PROSITE: PS00141; ASP_PROTENSE; UNKNOWN_1.
KW Hypothetical protein.
SQ
SEQUENCE 1491 AA; 170789 MW; 7DACBC98047FC409 CRC64;

Query Match
Best Local Similarity 85.4%; Score 35; DB 10; Length 1491;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGLVODG 8
DB 1035 GGLVODG 1042

RESULT 6
ID 006287 PRELIMINARY; PRT; 439 AA.
AC 006287;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE HYPOTHETICAL 37.1 KDA PROTEIN.
GN RV3595C OR MRCY07H7B.27.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxId=1773;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RA MEDLINE-98295987; PubMed-9634230; Garnier T., Churcher C., Harris D.,
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
DR EMBL: Z95557; CAB08933.1; -.
DR TUBERCULIST; RV3595C; -.
DR INTERPRO: IPR000084; -.
DR INTERPRO: IPR000087; -.
DR PFAM: PF000934; PE; 1.
DR PRODOM: PD001223; -.
KW Hypothetical protein.
SQ
SEQUENCE 439 AA; 37143 MW; B0BA83FC01AC560C CRC64;

Query Match
Best Local Similarity 82.9%; Score 34; DB 2; Length 439;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGLVODG 8
DB 302 GGLVODG 309

RESULT 7
ID 09LHDI PRELIMINARY; PRT; 1240 AA.
AC 09LHDI;

```

```

DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE MULTIDRUG RESISTANCE P-GLYCOPROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
RN 12
RN SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety p1,
RT TAC and BAC clones."
RL DNA Res. 7:217-221(2000).
DR EMBL: AP002051; BAB02627.1; -.
DR EMBL: AB026644; BAB02627.1; JOINED.
SQ SEQUENCE 1240 AA; 135640 MW; F00F17EBD4709173 CRC64;

Query Match
Best Local Similarity 82.9%; Score 34; DB 10; Length 1240;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGLVODG 8
DB 920 GGLVODG 927

RESULT 8
ID 061802 PRELIMINARY; PRT; 1510 AA.
AC 061802;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE H1E01.3 PROTEIN.
GN H1E01.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE-94150718; PubMed-7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copest T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Smaildon N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Spoat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Geisel C., Wamsley P.;
RL Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.
RN 13
RN SEQUENCE FROM N.A.

```

RC STRAIN-PRISTOL N2;
 RA Waterston R.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF067222; AAC17017.1;
 SQ SEQUENCE 1510 AA; 166889 MW; C1C08ACBCD6DFE11 CRC64;

Query Match
 Best Local Similarity 82.9%; Score 34; DB 5; Length 1510;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 GGLVODG 8
 Db 1495 GGLVODG 1502

RESULT 9
 OQ 092584 PRELIMINARY; PRT: 1542 AA.
 AD 092584;
 DT 01-MAY-1999 (TREMUREL: 10, Created)
 DT 01-MAY-1999 (TREMUREL: 10, Last sequence update)
 DT 01-OCT-2000 (TREMUREL: 15, Last annotation update)
 DE POLYPROTEIN.
 OS Lycopersicon esculentum (Tomato)
 CC Escherichia coli (strain K12) (E. coli) core eukaryotic; Asteridae; eumetazoa; I;
 CC Solanaceae; Solanum.
 NCBI_TaxID=4081;
 RX MEDLINE: 911251 PubMed=9926411;
 RA Parake M., Wulf B.B., Bonnem G., Thomas C.M., Jones D.A.,
 Jones J.D.;
 RT "Homologues of the Cf-9 disease resistance gene (Hcr9) are present at
 multiple loci on the short arm of tomato chromosome 1.";
 RL Moll. Plant Microbe Interact. 12:93-102(1999).
 DR INTERPRO: IPR000477;
 DR INTERPRO: IPR000953;
 DR INTERPRO: IPR001584;
 DR INTERPRO: IPR001969;
 DR PFAM: PF00078; ITC: 1;
 DR PROSITE: PS00651; ITC: 1;
 DR PROSITE: PS00113; CHROMO_2; 1.
 KW Polypeptide.
 SQ SEQUENCE 1542 AA; 175610 MW; OCC14A14A1455670 CRC64;

Query Match
 Best Local Similarity 92.9%; Score 34; DB 10; Length 1542;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GGLVODG 8
 Db 920 GGLVODG 927

RESULT 10
 ID P97157 PRELIMINARY; PRT: 303 AA.
 AC P97157;
 DT 01-MAY-1997 (TREMUREL: 03, Created)
 DT 01-MAY-1997 (TREMUREL: 03, Last sequence update)

DT 01-OCT-2000 (TREMUREL: 15, Last annotation update)
 DE 3 KDA PROTEIN.
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae.
 OC Escherichia.
 NCBI_TaxID=562;
 RX MEDLINE: 90158116 PubMed=2576094;
 RA Jajstakumari M.B., Thomas C.J., Halter R., Manning P.A.;
 RT "Genes for biosynthesis and assembly of CS3 pill of CFA/II
 enterotoxigenic Escherichia coli: novel regulation of pilus production
 by bypassing an amber codon.";
 RL Moll. Plant Microbe Interact. 12:93-102(1999).
 DR EMBL: X15944; CAA14818.1;
 DR INTERPRO: IPR000015;
 DR PFAM: PF00577; Usher: 1.
 SQ SEQUENCE 303 AA; 32839 MW; ABC62A5A60164542 CRC64;

Query Match
 Best Local Similarity 80.5%; Score 33; DB 2; Length 303;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 GGLVODG 8
 Db 103 GGLVODG 110

RESULT 11
 ID 064823 PRELIMINARY; PRT: 365 AA.
 AD 064823;
 DT 01-NOV-1996 (TREMUREL: 01, Created)
 DT 01-NOV-1996 (TREMUREL: 01, Last sequence update)
 DT 01-MAY-2000 (TREMUREL: 13, Last annotation update)
 DE FIBER PROTEIN.
 OS Human adenovirus type 37.
 CC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 NCBI_TaxID=36437;
 RX MEDLINE: 97167539 PubMed=9007080;
 RA Ardeberg N., Mel Y.F., Wedell G.;
 RT "Fiber genes of adenoviruses with tropism for the eye and the genital
 tract.";
 RL Virology 227:239-244(1997).
 DR INTERPRO: IPR000939;
 DR INTERPRO: IPR000978;
 DR PFAM: PF00541; adeno_fiber: 1.
 DR PFAM: PF00608; adeno_fiber2: 2.
 DR PROSITE: PS00107; ADENOVIRFIBRE.
 KW Fiber protein.
 SQ SEQUENCE 365 AA; 40092 MW; 2F713268D55AD80 CRC64.

Query Match
 Best Local Similarity 80.5%; Score 33; DB 14; Length 365;

```

Best Local Similarity 75.0%; Pred. No. 96;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGLVQDG 8
   |||: |||
Db 67 GGLTQDG 74

RESULT 12
Q64822 PRELIMINARY: PRT: 365 AA.
ID 064822
AC 064822:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE FIBER PROTEIN.
OS Human adenovirus type 19.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=28278;
RN [1]
RP SEQUENCE FROM N.A.
RA Pring-Akerblom P., Helm A., Trisjsenaar F.E.J.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=97167539; PubMed=9007080;
RX Arberg N., Mei Y.F., Wadell G.;
RA Arberg N., Mei Y.F., Wadell G.;
RT "Fiber genes of adenoviruses with tropism for the eye and the genital tract."
RL Virology 227:239-244(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=587;
RA Arberg N., Mei Y.F., Wadell G.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: X94485; CA64231.1; -;
DR EMBL: U69130; AAB7132.1; -;
DR HSP: P11818; 1KB.
DR INTERPRO: IPR000931; -;
DR INTERPRO: IPR000939; -;
DR INTERPRO: IPR000978; -;
DR PFAM: PF00541; adeno_fiber.1;
DR PFAM: PF00608; adeno_fiber2.2;
DR PRINTS: PR00307; ADENOVSFIBRE.
KW Fiber protein.
SQ SEQUENCE 365 AA; 40131 MW; E7B99E50EBB38FAF CRC64;

Query Match 80.5%; Score 33; DB 14; Length 365;
Best Local Similarity 75.0%; Pred. No. 96;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGLVQDG 8
   |||: |||
Db 67 GGLTQDG 74

RESULT 13
Q04927 PRELIMINARY: PRT: 476 AA.
ID 004927
AC 004927:
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ALTIMASE (EC 4.4.1.4) (ALTIMIN LYASE) (CYSTEINE SULPHOXIDE LYASE).
OS Allium tuberosum.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Asparagales; Alliaceae; Allium.
OX NCBI_TaxID=4683;
RN [1]
RP SEQUENCE FROM N.A.
RA Manabe T., Saito K.;

```

```

RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: AN S-ALKYL-L-CYSTEINE S-OXIDE - AN ALKYL
CC SULFENATE + 2 AMINOCRYLATE.
CC -I- COFACTOR: PYRIDOXAL-PHOSPHATE.
DR EMBL: AB004269; BAA20358.1; -;
DR MENDEL: 16931; Alltu:1125;16931.
DR INTERPRO: IPR000561; -;
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PRODOM: PD014435; -; 1.
SQ SEQUENCE 476 AA; 54084 MW; 074B696F559B6492 CRC64;

Query Match 80.5%; Score 33; DB 10; Length 476;
Best Local Similarity 62.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGLVQDG 8
   |||: |||
Db 413 GGLTQDG 420

RESULT 14
Q53884 PRELIMINARY: PRT: 606 AA.
ID 053884
AC 053884:
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE PGRS-FAMILY PROTEIN.
GN RV0872C OR MTY043.65C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634220;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sultson J.E.,
RA Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence."
RL Nature 393:537-544(1998).
DR EMBL: AL022004; CAAT678.1; -;
DR TUBERCULIST: RV0872C; -;
DR INTERPRO: IPR000084; -;
DR PFAM: PF00934; PE.1.
DR PRODOM: PD001223; -; 1.
SQ SEQUENCE 606 AA; 50268 MW; 4AFA2DDB6A1EF10 CRC64;

Query Match 80.5%; Score 33; DB 2; Length 606;
Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGLVQDG 8
   |||: |||
Db 397 GGLTQDG 404

RESULT 15
Q9RW35 PRELIMINARY: PRT: 213 AA.
ID 09RW35
AC 09RW35:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

```

DT 01-JUN-2000 (TRENBLUREL, 14, last annotation update)
 DE NITROGEN REGULATOR, PUTATIVE.
 GN DR0834.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-R1;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome Sequence of the Radioresistant Bacterium Deinococcus
 RT radiodurans R1.";
 RL Science 286:1571-1577(1999).
 DR EMBL; AE001937; AAF10412.1; -.
 DR TIGR; DR0834; -.
 DR INTERPRO; IPR000595; -.
 DR INTERPRO; IPR001808; -.
 DR PFAM; PF00027; CNMP_binding; 1.
 DR PFAM; PF00325; crp; 1.
 DR PRINTS; PR00034; HTHCRP.
 DR PROSITE; PS50042; CNMP_BINDING-3; 1.
 DR PROSITE; PS50042; CNMP_BINDING-3; 1.
 SQ SEQUENCE 213 AA; 23266 MW; A8200332A7C252D7 CRC64;

Query Match 78.0%; Score 32; DB 2; Length 213;
 Best Local Similarity 75.0%; Pred. No. 82;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGLVQDG 8
 ||: ||||
 Db 121 GGIHVQDG 128

Search completed: June 13, 2001, 14:20:37
 Job time: 737 sec.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:16:37 ; Search time 118.55 Seconds
(without alignments)
1.296 Million cell updates/sec

Title: PCT-US01-05825A-20

Perfect score: 41

Sequence: 1 GGLVQDG 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents, AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	80.5	115	2	US-07-903-029-4
2	33	80.5	115	2	US-07-903-029-5
3	32	78.0	346	1	US-08-774-169-3
4	31	75.6	20	2	US-08-859-931A-4
5	30	73.2	247	2	US-08-851-974-4
6	30	73.2	247	2	US-09-213-390-4
7	30	73.2	368	1	US-08-093-372-4
8	30	73.2	492	3	US-09-006-636-4
9	30	73.2	492	4	US-09-006-632-4
10	29.5	72.0	580	3	US-08-482-677-10
11	29	70.7	118	3	US-08-545-809A-108
12	29	70.7	123	2	US-08-428-197-38
13	29	70.7	123	5	PCR-US93-10555-38
14	29	70.7	220	4	US-09-004-731-95
15	29	70.7	220	4	US-08-749-699-95
16	29	70.7	226	4	US-09-004-731-8
17	29	70.7	226	4	US-08-749-699-8
18	29	70.7	254	4	US-09-004-731-89
19	29	70.7	254	4	US-09-004-731-92
20	29	70.7	254	4	US-08-749-699-89
21	29	70.7	254	4	US-08-749-699-92
22	29	70.7	327	4	US-09-004-731-2
23	29	70.7	327	4	US-09-004-731-5
24	29	70.7	327	4	US-08-749-699-2
25	29	70.7	327	4	US-08-749-699-5
26	29	70.7	375	1	US-07-817-917-1
27	29	70.7	401	2	US-08-741-327E-11

ALIGNMENTS

28	29	70.7	511	1	US-08-220-151-17	Sequence 17, Appl
29	29	70.7	511	1	US-08-413-118-17	Sequence 17, Appl
30	29	70.7	511	3	US-08-473-446-17	Sequence 17, Appl
31	29	70.7	552	3	US-09-120-365-5	Sequence 5, Appl
32	29	70.7	674	1	US-08-317-522A-3	Sequence 3, Appl
33	29	70.7	674	1	US-08-439-818A-3	Sequence 3, Appl
34	29	70.7	674	2	US-08-751-965-3	Sequence 3, Appl
35	29	70.7	674	2	US-08-738-975-3	Sequence 3, Appl
36	29	70.7	674	2	US-08-728-626-3	Sequence 3, Appl
37	29	70.7	674	4	US-08-808-599A-3	Sequence 3, Appl
38	29	70.7	749	1	US-08-317-522A-2	Sequence 2, Appl
39	29	70.7	749	1	US-08-439-818A-2	Sequence 2, Appl
40	29	70.7	749	2	US-08-751-965-2	Sequence 2, Appl
41	29	70.7	749	2	US-08-738-975-2	Sequence 2, Appl
42	29	70.7	749	2	US-08-728-626-2	Sequence 2, Appl
43	29	70.7	749	4	US-08-808-599A-2	Sequence 2, Appl
44	29	70.7	1408	1	US-08-612-521-2	Sequence 2, Appl
45	28	68.3	27	2	US-08-709-874A-1	Sequence 1, Appl

RESULT 1
US-07-903-029-4
Sequence 4, Application US/07903029
Patent No. 5969097
GENERAL INFORMATION:
APPLICANT: Wiegand, Roger C.
APPLICANT: Currie, Mark C.
APPLICANT: Fok, Kam F.
TITLE OF INVENTION: Human Guanylin
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis A. Bennett, Monsanto Co., A35G
STREET: 800 N. Lindbergh Blvd.
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63167
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/903,029
FILING DATE: 19920623
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Dennis A.
REGISTRATION NUMBER: 34,547
REFERENCE/DOCKET NUMBER: 07-21(872)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)694-5402
TELEFAX: (314)694-9009
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 115 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-903-029-4

Query Match 80.5%; Score 33; DB 2; length 115;
Best Local Similarity 75.0%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGLVQDG 8
11:1111
DB 20 GGLVQDG 27

```
RESULT 2
US-07-903-029-5
; Sequence 5, Application US/07903029
; Patent No. 5969097
; GENERAL INFORMATION:
; APPLICANT: Wiegand, Roger C.
; APPLICANT: Currie, Mark C.
; APPLICANT: Fok, Kam F.
; TITLE OF INVENTION: Human Guanylin
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, Monsanto Co., A3SG
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/903,029
; FILING DATE: 19920623
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A.
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: 07-21(872)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)694-5402
; TELEFAX: (314)694-9009
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 115 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-903-029-5

Query Match      80.5%; Score 33; DB 2; Length 115;
Best Local Similarity 75.0%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGLVODG 8
   ||:||||
Db 20 GGVTVDG 27

RESULT 3
US-08-774-169-3
; Sequence 3, Application US/08774169
; Patent No. 5756332
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: A NOVEL GUANOSINE MONOPHOSPHATE REDUCTASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
```

```
APPLICATION NUMBER: US/08/774,169
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0182 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 346 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 473772
US-08-774-169-3

Query Match      78.0%; Score 32; DB 1; Length 346;
Best Local Similarity 62.5%; Pred. No. 85;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGLVODG 8
   ||:||||
Db 213 GGMIVSDG 220

RESULT 4
US-08-859-931A-4
; Sequence 4, Application US/08859931A
; Patent No. 5945510
; GENERAL INFORMATION:
; APPLICANT: PASANO, Alessio
; TITLE OF INVENTION: SUBSTANTIALLY PURE ZONULIN, A
; TITLE OF INVENTION: PHYSIOLOGICAL MODULATOR OF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W., Suite 800
; CITY: Washington, D.C.
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/859,931A
; FILING DATE: 21 MAY 1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: KIP, Gordon
; REGISTRATION NUMBER: 30,764
; REFERENCE/DOCKET NUMBER: A-6901
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```

MOLECULE TYPE: Peptide
HYPOTHETICAL: NO
US-08-859-931A-4

Query Match 75.6%; Score 31; DB 2; Length 20;
Best Local Similarity 75.0%; Pred. No. 6.7;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GGLVQDG 8
111111
Db 8 GGVLVQPG 15

RESULT 5
US-08-851-974-4
Sequence 4, Application US/08851974
Patent No. 5858758
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN SERINE PROTEASE PRECURSOR
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851.974
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0288 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-845-4166
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 247 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 306682
US-08-851-974-4

Query Match 73.2%; Score 30; DB 2; Length 247;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GGLVQDG 7
111111
Db 50 GGFLIQD 56

RESULT 6
US-09-213-390-4
Sequence 4, Application US/09213390

Patent No. 5965711
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN SERINE PROTEASE PRECURSOR
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/213.390
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/851.974
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0288 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-845-4166
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 247 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 306682
US-09-213-390-4

Query Match 73.2%; Score 30; DB 2; Length 247;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GGLVQDG 7
111111
Db 50 GGFLIQD 56

RESULT 7
US-08-093-372-4
Sequence 4, Application US/08093372
Patent No. 5530187
GENERAL INFORMATION:
APPLICANT: Lamb, Christopher J.
APPLICANT: Zhu, Qun
APPLICANT: Maher, Eileen A.
APPLICANT: Dixon, Richard A.
TITLE OF INVENTION: TRANSGENIC PLANTS CONTAINING MULTIPLE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: U.S.A.
ZIP: 90071-2921
COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/093,372
FILING DATE: 16-JUL-1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-4737
TELEFAX: 619-546-9392
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 368 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-093-372-4

Query Match 73.2%; Score 30; DB 1; Length 368;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 GGLVQDG 8
|:|||||
Db 234 GVMVQDG 240

RESULT 8
US-09-006-636-4
Sequence 4, Application US/09006636
Patent No. 6005092
GENERAL INFORMATION:
APPLICANT: Shoseyov, Oded
TITLE OF INVENTION: ARABIDOPSIS THALIANA ENDO-1,4-BETA-
TITLE OF INVENTION: GLUCANASE GENE, PROMOTER AND PROTEIN
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/006,636
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Baldwin, Geraldine F.
REGISTRATION NUMBER: 31,232
REFERENCE/DOCKET NUMBER: 7809-019
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 492 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
US-09-006-636-4

Query Match 73.2%; Score 30; DB 3; Length 492;
Best Local Similarity 75.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGLVQDG 8
|:|||||
Db 328 GGLVKTG 335

RESULT 9
US-09-006-632-4
Sequence 4, Application US/09006632
Patent No. 6184440
GENERAL INFORMATION:
APPLICANT: Shoseyov, Oded
APPLICANT: Shigel, Etai
TITLE OF INVENTION: TRANSGENIC PLANTS OF ALTERED
TITLE OF INVENTION: MORPHOLOGY
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/006,632
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Baldwin, Geraldine F.
REGISTRATION NUMBER: 31,232
REFERENCE/DOCKET NUMBER: 7809-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 492 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-006-632-4

Query Match 73.2%; Score 30; DB 4; Length 492;
Best Local Similarity 75.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGLVQDG 8
|:|||||
Db 328 GGLVKTG 335

RESULT 10
US-08-482-677-10
Sequence 10, Application US/08482677
Patent No. 6017714
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc

APPLICANT: Serafini, Tito
APPLICANT: Kennedy, Timothy
APPLICANT: Placzek, Marysia
APPLICANT: Jessel, Thomas
APPLICANT: Dodd, Jane
TITLE OF INVENTION: Netrlins
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,677
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: 0093-300-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 580 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-482-677-10

Query Match 72.0%; Score 29.5; DB 3; Length 580;
Best Local Similarity 72.7%; Pred. No. 4.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Oy 1 GGLV---VQDG 8
1111 1111
Db 188 GGLAFSVQDG 198

RESULT 11
US-08-545-809A-108
Sequence 108, Application US/08545809A
Patent No. 6096878
GENERAL INFORMATION:
APPLICANT: Honjo, Tasaku
APPLICANT: Matsuda, Rumiiko
TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,809A

FILING DATE: 27-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 06501/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 108:
SEQUENCE CHARACTERISTICS:
LENGTH: 118 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-545-809A-108

Query Match 70.7%; Score 29; DB 3; Length 118;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GGLVQDG 8
1111 1
Db 27 GGLVQDG 34

RESULT 12
US-08-428-197-38
Sequence 38, Application US/08428197
Patent No. 5891438
GENERAL INFORMATION:
APPLICANT: SILVERMAN, GREGG J.
TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF
TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED
TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERANTIGEN AND CONJUGATES
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Spensley Horn Judas & Lubitz
STREET: 1880 Century Park East - Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/428,197
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10555
FILING DATE: 29-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Howells, Stacy L.
REGISTRATION NUMBER: 34,842
REFERENCE/DOCKET NUMBER: FD-2630
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-428-197-38

Query Match 70.7%; Score 29; DB 2; Length 123;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGLVODG 8
11111
DB 8 GGLVQPG 15

RESULT 13

PCT-US93-10555-38
Sequence 38, Application PC/TUS9310555
GENERAL INFORMATION:

APPLICANT: SILVERMAN, GREGG J.
TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF
TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH
TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERANTIGEN AND CONJUGATES
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Spensley Horn Jubas & Lubitz
STREET: 1880 Century Park East - Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10555
FILING DATE: 29-OCT-1993
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Howells, Stacy L.
REGISTRATION NUMBER: 34,842
REFERENCE/DOCKET NUMBER: FD-2630
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO.: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-10555-38

Query Match 70.7%; Score 29; DB 5; Length 123;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGLVODG 8
11111
DB 8 GGLVQPG 15

RESULT 14

US-09-004-731-95
Sequence 95, Application US/09004731
Patent No. 6177258
GENERAL INFORMATION:

APPLICANT: Wu Hunter, Shirley
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
TITLE OF INVENTION: MOLECULES AND USES THEREOF
ATTORNEY/AGENT INFORMATION:

NUMBER OF SEQUENCES: 103

CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/004,731
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/749,699
FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO.: 95:
SEQUENCE CHARACTERISTICS:
LENGTH: 220 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-004-731-95

Query Match 70.7%; Score 29; DB 4; Length 220;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGLVODG 8
11111
DB 91 GGLVQDG 98

RESULT 15
US-08-749-699-95
Sequence 95, Application US/08749699
Patent No. 6210920
GENERAL INFORMATION:

APPLICANT: Wu Hunter, Shirley
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
TITLE OF INVENTION: MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/749,699
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ. ID NO.: 95:
SEQUENCE CHARACTERISTICS:
LENGTH: 220 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-749-639-95

Query Match 70.7%; Score 29; DB 4; Length 220;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGLVQDG 8
|||
Db 91 GGYLGQDG 98

Search completed: June 13, 2001, 14:16:37
JOB Time: 498 sec

Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:14:41 ; Search time 150.28 Seconds
(without alignments)
3.043 Million cell updates/sec

Title: PCT-US01-05825A-21

Perfect score: 49
Sequence: 1 GCGCVQPE 8

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: A.Geneseq_0401:*
2: /SID6/gcgdata/geneseq/geneseq/AA1980.DAT:*
3: /SID6/gcgdata/geneseq/geneseq/AA1981.DAT:*
4: /SID6/gcgdata/geneseq/geneseq/AA1982.DAT:*
5: /SID6/gcgdata/geneseq/geneseq/AA1983.DAT:*
6: /SID6/gcgdata/geneseq/geneseq/AA1984.DAT:*
7: /SID6/gcgdata/geneseq/geneseq/AA1985.DAT:*
8: /SID6/gcgdata/geneseq/geneseq/AA1986.DAT:*
9: /SID6/gcgdata/geneseq/geneseq/AA1987.DAT:*
10: /SID6/gcgdata/geneseq/geneseq/AA1988.DAT:*
11: /SID6/gcgdata/geneseq/geneseq/AA1989.DAT:*
12: /SID6/gcgdata/geneseq/geneseq/AA1990.DAT:*
13: /SID6/gcgdata/geneseq/geneseq/AA1991.DAT:*
14: /SID6/gcgdata/geneseq/geneseq/AA1992.DAT:*
15: /SID6/gcgdata/geneseq/geneseq/AA1993.DAT:*
16: /SID6/gcgdata/geneseq/geneseq/AA1994.DAT:*
17: /SID6/gcgdata/geneseq/geneseq/AA1995.DAT:*
18: /SID6/gcgdata/geneseq/geneseq/AA1996.DAT:*
19: /SID6/gcgdata/geneseq/geneseq/AA1997.DAT:*
20: /SID6/gcgdata/geneseq/geneseq/AA1998.DAT:*
21: /SID6/gcgdata/geneseq/geneseq/AA1999.DAT:*
22: /SID6/gcgdata/geneseq/geneseq/AA2000.DAT:*
23: /SID6/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	49	100.0	8	21	V79125	Peptide antagonist
2	43	87.8	31	20	W88384	Human Zneul BGF-11
3	43	87.8	73	20	W88389	Human Zneul partia
4	43	87.8	115	21	B41718	Human Zneul ORF1482
5	43	87.8	153	21	B41638	Human ORFX ORF1402
6	43	87.8	158	20	W88388	Human Zneul partia
7	43	87.8	169	20	W88390	Human Zneul partia
8	43	87.8	224	20	W88370	Human normal uteru
9	43	87.8	247	21	Y52139	Human TANGO 125b (
10	43	87.8	254	20	W88382	Human neuro-growth
11	43	87.8	265	21	B42204	Human ORFX ORF1968

12	43	87.8	273	20	Y41769	Human PRO213-1 pro
13	43	87.8	273	20	Y41770	Human PRO1330 prot
14	43	87.8	273	20	Y41771	Human PRO1449 prot
15	43	87.8	273	20	W88381	Human neuro-growth
16	43	87.8	273	21	B44325	Human PRO213-1 pro
17	43	87.8	273	21	B44326	Human PRO1330 prot
18	43	87.8	273	21	B44327	Human PRO1449 prot
19	43	87.8	273	21	B18673	Amino acid sequenc
20	43	87.8	273	21	B18674	Amino acid sequenc
21	43	87.8	273	21	B18675	Amino acid sequenc
22	43	87.8	273	21	B24042	Human PRO213 prote
23	43	87.8	273	21	B24043	Human PRO1330 prot
24	43	87.8	273	21	B24044	Human PRO1449 prot
25	43	87.8	273	21	B01376	Neuron-associated
26	43	87.8	273	21	Y52137	Human TANGO 125 (T
27	43	87.8	295	20	Y41685	Human PRO213 prote
28	43	87.8	295	21	B44241	Human PRO213 (UNO1
29	43	87.8	307	21	B41644	Human ORFX ORF1408
30	41	83.7	8	21	Y79113	Peptide antagonist
31	41	83.7	8	21	Y79113	Peptide antagonist
32	40	81.6	8	21	Y79116	Peptide antagonist
33	39	79.6	8	21	Y79117	Peptide antagonist
34	39	79.6	8	21	Y79121	Peptide antagonist
35	39	79.6	13	20	W86096	Peptide from human
36	39	79.6	13	20	W86096	Peptide from human
37	39	79.6	13	20	W86096	Peptide from human
38	39	79.6	18	16	W73953	Peptide from human
39	39	79.6	18	16	W73953	IgG heavy chain fr
40	39	79.6	20	20	W94487	N-terminal of C179
41	39	79.6	20	21	Y79130	Human adult heart
42	39	79.6	26	15	R49514	Human adult heart
43	39	79.6	30	17	R97323	Human adult heart
44	39	79.6	30	17	R97324	Human adult heart
45	39	79.6	30	17	R87049	Human adult heart

ALIGNMENTS

RESULT 1	Y79125	standard; peptide; 8 AA.
ID	Y79125	
XX	Y79125	
AC	Y79125	
XX	05-JUN-2000	(first entry)
XX		
DE	Peptide antagonist of zonulin.	
XX		
KW	zonulin; antagonist; zonula occludens toxin receptor;	
KW	blood-brain barrier; antiinflammatory; cerebroprotective;	
KW	neuroprotective; dermatological; antidiabetic; antiviral;	
KW	antibacterial; cytostatic; anti-HIV; vulnirary; antiallergic;	
KW	hypocensative; immunosuppressive; antiparasitic; vasotropic;	
KW	gastrointestinal inflammation; therapy.	
XX		
OS	Synthetic.	
XX		
PN	WO200007609-A1.	
XX		
PD	17-FEB-2000.	
XX		
PF	28-JUL-1999;	99WO-US16683.
XX		
PR	03-AUG-1998;	98US-0127815.
XX		
PA	(UYMA-) UNIV MARYLAND BALTIMORE.	
XX		
PI	Fasano A.	
XX		
DR	WPI; 2000-205565/18.	
XX		
PT	New peptide antagonist of zonulin useful as antiinflammatory agent for	

PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 XX
 PS Claim 1; Page 47; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see 479105-29) of the invention
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcooidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 CC
 SQ Sequence 8 AA:

Query Match 100.0%; Score 49; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCVQPG 8
 |||||
 Db 1 ggcvcvpg 8

RESULT 2
 W88384
 ID W88384 standard; Peptide; 31 AA.

AC W88384;

DT 26-APR-1999 (first entry)

DE Human Zneul EGF-like domain 1.

KW Zneul-1; neuro-growth factor-like protein; human; breast cancer;
 KW glioblastoma; pituitary adenoma; Alzheimer's disease; therapy;
 KW nerve regeneration; haematopoiesis; fertility; contraception;
 KW antibody; epidermal growth factor; EGF.

OS Homo sapiens.

PN W09857983-A2.

PD 23-DEC-1998.

PF 18-JUN-1998; 98WO-US12763.

PR 18-JUN-1997; 97US-0878322.

PR 18-JUN-1997; 97US-0050143.

PA (ZYMO) ZYMOGENETICS INC.

PI Blumberg H, Jelinek LJ, Lehner JM, Sheppard PO;

PI Whitmore TE;

XX WPI; 1999-095324/08.

PT New mammalian Zneul polypeptides - used to, e.g. treat Alzheimer's
 PT disease, cancer and to repopulate blood cells

PS Claim 6; Page 53; 70pp; English.

CC This peptide comprises the first epidermal growth factor-like
 CC domain (EGF-1), i.e. amino acid residues 105-135, of human Zneul
 CC (see W88381), a new neuro-growth factor-like protein. Zneul can be
 CC used as a growth, maintenance, or differentiation factor in the
 CC spinal cord, heart, spleen, testis, thyroid and lymph nodes. It
 CC may also play a role in breast cancer, glioblastomas, and pituitary
 CC adenomas. Zneul may also be used to treat Alzheimer's disease,
 CC cancer, to repopulate blood cells after chemotherapy, to stimulate
 CC myofibroblast proliferation, stimulate or inhibit growth factors
 CC made in the placenta, in fertility and contraception, or to
 CC regenerate nerves. Claimed Zneul polypeptides (see also W88382-97),
 CC including specific domains of Zneul and epitope-bearing portions of
 CC Zneul, can be used to raise specific antibodies for use e.g. in
 CC diagnostic assays.

SQ Sequence 31 AA:

Query Match 87.8%; Score 43; DB 20; Length 31;
 Best Local Similarity 87.5%; Pred. No. 3.5;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCVQPG 8
 |||||
 Db 10 ggcvcvpg 17

RESULT 3
 W88389
 ID W88389 standard; Protein; 73 AA.

AC W88389;

DT 26-APR-1999 (first entry)

DE Human Zneul partial polypeptide.

KW Zneul-1; neuro-growth factor-like protein; human; breast cancer;
 KW glioblastoma; pituitary adenoma; Alzheimer's disease; therapy;
 KW nerve regeneration; haematopoiesis; fertility; contraception;
 KW antibody.

OS Homo sapiens.

PN W09857983-A2.

PD 23-DEC-1998.

PF 18-JUN-1998; 98WO-US12763.

PR 18-JUN-1997; 97US-0878322.

PR 18-JUN-1997; 97US-0050143.

PA (ZYMO) ZYMOGENETICS INC.

PI Blumberg H, Jelinek LJ, Lehner JM, Sheppard PO;

PI Whitmore TE;

DR WPI; 1999-095324/08.

PT New mammalian Zneul polypeptides - used to, e.g. treat Alzheimer's
 PT disease, cancer and to repopulate blood cells

PS Claim 6; Page 57; 70pp; English.

XX This claimed polypeptide comprises amino acid residues 105-177 of
CC the human Zneul mature protein (see W88382), i.e. the 2 epidermal
CC growth factor-like domains (see also W88384-85) of Zneul. Zneul is
CC a new neuro-growth factor-like protein that can be used as a
CC growth, maintenance, or differentiation factor in the spinal cord,
CC heart, spleen, testis, thyroid and lymph nodes. Zneul may also
CC play a role in breast cancer, glioblastomas, and pituitary
CC adenomas. Zneul can be used to treat Alzheimer's disease, cancer,
CC to repopulate blood cells after chemotherapy, to stimulate
CC myofibroblast proliferation, stimulate or inhibit growth factors
CC made in the placenta, in fertility and contraception, or to
CC regenerate nerves. Claimed Zneul polypeptides (see also W88387-97),
CC including specific domains of Zneul and epitope-bearing portions of
CC Zneul, can be used to raise specific antibodies for use e.g. in
CC diagnostic assays.
XX
SQ Sequence 73 AA:

Query Match 87.8% Score 43; DB 20; Length 73;
Best Local Similarity 87.5% Pred. No. 7.4;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 GGCVCVPC 8
111111
Db 10 ggcvcvpg 17

RESULT 4
B41718
ID B41718 standard; Protein; 115 AA.
XX
AC B41718;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF1482 polypeptide sequence SEQ ID NO:2964.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antipsoptic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antineumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
OS Homo sapiens.
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PE 31-MAR-2000; 2000WO-US08621.
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Shinkets RA, Teach M;
XX
XX WPI: 2000-602362/57.
XX
XX DR N-PSDB; C75927.
XX

XX Novel nucleic acids and peptides derived from open reading frame X.
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 11; Page 2200; 5507pp; English.
XX
CC C74446 to C77606 encode the proteins given in B40237 to B43397, which
CC represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antipsoptic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antineumatic; antithyroid; and antianaemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 115 AA:

Query Match 87.8% Score 43; DB 21; Length 115;
Best Local Similarity 87.5% Pred. No. 11;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 GGCVCVPC 8
111111
Db 19 ggcvcvpg 26

RESULT 5
B41638
ID B41638 standard; Protein; 153 AA.
XX
AC B41638;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF1402 polypeptide sequence SEQ ID NO:2804.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antipsoptic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antineumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
OS Homo sapiens.
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PE 31-MAR-2000; 2000WO-US08621.
XX
XX

PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR WPI: 2000-602362/57.
 DR N-PSDB; C75847.
 XX

PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11; Page 204; 5507pp; English.

CC C7446 to C77606 encode the proteins given in B40237 to B4397, which
 CC represent the human ORF open reading frames 1 to 3161. The ORF
 CC sequences have activities such as: cytosolic; hepatotropic; vulnary;
 CC antiproliferative; antiparkinsonian; neuroprotective; osteoprotective;
 CC anticonvulsant; antidiabetic; immunosuppressive; immunostimulant;
 CC dermatological; coagulant; vasotrophic; antidiabetic; hypotensive;
 CC antiviral; antifungal; antineoplastic; antitumor; antidiabetic;
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORF-associated disorder. The nucleic acids can be used to express ORF
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, anti-inflammatory disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 153 AA;

Query Match 87.8%; Score 43; DB 21; Length 153;
 Best Local Similarity 87.5%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCVCVPG 8
 Db 91 ggcvcvpg 98

RESULT 6
 W88388
 ID W88388 standard; Protein; 158 AA.
 AC W88388;
 XX
 DT 26-APR-1999 (first entry)
 XX
 DE Human Zneu1 partial polypeptide.

XX Zneu-1; neuro-growth factor-like protein; human; breast cancer;
 KW glioblastoma; pituitary adenoma; Alzheimer's disease; therapy;
 KW nerve regeneration; haematopoiesis; fertility; contraception;
 KW antibody.
 XX
 OS Homo sapiens.
 XX
 PN W09857983-A2.
 XX
 PD 23-DEC-1998.
 XX
 PF 18-JUN-1998; 98MO-US12763.

XX
 PR 18-JUN-1997; 97US-0878322.
 PR 18-JUN-1997; 97US-0050143.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Blumberg H, Jelinek LJ, Lehner JM, Sheppard FO;
 XX
 DR WPI: 1999-095324/08.
 DR
 XX
 PT New mammalian Zneu1 polypeptides - used to, e.g. treat Alzheimer's
 PT disease, cancer and to repopulate blood cells
 XX
 PS Claim 6; Page 56; 70pp; English.

CC This claimed polypeptide comprises amino acid residues 1-158 of the
 CC human Zneu1 mature protein (see W88382), i.e. the first HSMC3W5A
 CC homology domain and first EGF-like domain of Zneu1. Zneu1 is a new
 CC neuro-growth factor-like protein that can be used as a growth,
 CC maintenance, or differentiation factor in the spinal cord, heart,
 CC spleen, testis, thyroid and lymph nodes. Zneu1 may also play a
 CC role in breast cancer, glioblastomas, and pituitary adenomas.
 CC Zneu1 can be used to treat Alzheimer's disease, cancer, to
 CC repopulate blood cells after chemotherapy, to stimulate
 CC myofibroblast proliferation, stimulate or inhibit growth factors
 CC made in the placenta, in fertility and contraception, or to
 CC regenerate nerves. Claimed Zneu1 polypeptides (see also W88382-97),
 CC including specific domains of Zneu1 and epitope-bearing portions of
 CC Zneu1, can be used to raise specific antibodies for use e.g. in
 XX
 SQ Sequence 158 AA;

Query Match 87.8%; Score 43; DB 20; Length 158;
 Best Local Similarity 87.5%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCVCVPG 8
 Db 95 ggcvcvpg 102

RESULT 7
 W88390
 ID W88390 standard; Protein; 169 AA.
 AC W88390;
 XX
 DT 26-APR-1999 (first entry)
 XX
 DE Human Zneu1 partial polypeptide.

XX Zneu-1; neuro-growth factor-like protein; human; breast cancer;
 KW glioblastoma; pituitary adenoma; Alzheimer's disease; therapy;
 KW nerve regeneration; haematopoiesis; fertility; contraception;
 KW antibody.
 XX
 OS Homo sapiens.
 XX
 PN W09857983-A2.
 XX
 PD 23-DEC-1998.
 XX
 PF 18-JUN-1998; 98MO-US12763.
 XX
 PR 18-JUN-1997; 97US-0878322.
 PR 18-JUN-1997; 97US-0050143.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Blumberg H, Jelinek LJ, Lehner JM, Sheppard FO;

PI Whitmore TE;
 XX
 DR WPI: 1999-095324/08.
 XX
 PT New mammalian Zneul polypeptides - used to, e.g. treat Alzheimer's
 XX disease, cancer and to repopulate blood cells
 XX
 PS Claim 6; Page 57-58; 70pp; English.
 XX
 CC This claimed polypeptide comprises amino acid residues 105-273 of
 CC the human Zneul mature protein (see W88382), i.e. the 2 epidermal
 CC growth factor-like domains (see also W88384-85) and the second
 CC HSMC35A homology domain of Zneul. Zneul is a new neuro-growth
 CC factor-like protein that can be used as a growth, maintenance, or
 CC differentiation factor in the spinal cord, heart, spleen, testis,
 CC thyroid and lymph nodes. Zneul may also play a role in breast
 CC cancer, glioblastomas, and pituitary adenomas. Zneul can be used
 CC to treat Alzheimer's disease, cancer, to repopulate blood cells
 CC after chemotherapy, to stimulate myofibroblast proliferation,
 CC stimulate or inhibit growth factors made in the placenta, in
 CC fertility and contraception, or to regenerate nerves. Claimed
 CC Zneul polypeptides (see also W88382-97), including specific domains
 CC of Zneul and epitope-bearing portions of Zneul, can be used to
 CC raise specific antibodies for use e.g. in diagnostic assays.
 CC
 SQ Sequence 169 AA:

Query Match 87.8%; Score 43; DB 20; Length 169;
 Best Local Similarity 87.5%; Pred. No. 15;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCVQPG 8
 II IIIII
 DB 10 ggcvcvpg 17

RESULT 8
 Y59870 1999-095324/08.
 XX Y59870 standard; Protein: 224 AA.
 AC Y59870;
 XX
 DT 19-JAN-2000 (first entry)
 XX
 DE Human normal uterus tissue derived protein 33.
 XX
 KM Human; uterus; cancer; treatment; anticancer; cytostatic; gene therapy;
 XX EST; expressed sequence tag.
 OS Homo sapiens.
 XX
 PN DE19817946-A1.
 XX
 PD 21-OCT-1999.
 XX
 PF 17-APR-1998; 98DE-1017946.
 XX
 PR 17-APR-1998; 98DE-1017946.
 XX
 PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX
 PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
 XX
 DR WPI: 1999-591956/51.
 DR N-PSDB: Z41336.
 XX
 PT New nucleic acid sequences expressed in normal uterine tissues, and
 PT derived polypeptides, for treatment of uterine cancer and
 PT identification of therapeutic agents
 PS Claim 23; Page 134; 154pp; German.
 XX

CC This invention describes novel cDNA sequences (A) highly expressed in
 CC normal uterine tissue which can have anticancer and cytostatic activity
 CC and can be used for gene therapy. (A) are used (i) for recombinant
 CC expression of polypeptides (B) and (ii) to isolate complete genes.
 CC (B) are used (i) to identify agents suitable for treatment of uterine
 CC cancer; (ii) directly for treating this form of cancer (including
 CC expression from gene therapy vectors) and (iii) for generation of
 CC specific antibodies. (A) are identified by assembling ESTs (expressed
 CC sequence tags) from a particular tissue type before comparison of
 CC expression patterns. This allows a significantly longer fragment of the
 CC gene to be revealed, so should reduce the number of failures associated
 CC with the fact that ESTs from different libraries may represent
 CC different parts of the same unknown gene, distorting the estimated
 CC frequency of occurrence in a particular tissue. Y59838-Y59892 represent
 CC protein fragments encoded by the human uterine tissue derived cDNA
 CC fragments represented in Z41325-Z41385.
 CC
 SQ Sequence 224 AA:

Query Match 87.8%; Score 43; DB 20; Length 224;
 Best Local Similarity 87.5%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCVQPG 8
 II IIIII
 DB 65 ggcvcvpg 72

RESULT 9
 Y52139 1999-095324/08.
 XX Y52139 standard; Protein: 247 AA.
 AC Y52139;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human TANGO 125b (T125b) amino acid sequence.
 XX
 KM TANGO 125; T125; alternative splice variant; EGF domain; antibody;
 KM secreted protein; agonist; antagonist; predictive medicine; treatment;
 XX forensic biology.
 OS Homo sapiens.
 XX
 PN WO954437-A2.
 XX
 PD 28-OCT-1999.
 XX
 PF 23-APR-1999; 99WO-US08900.
 XX
 PR 23-APR-1998; 98US-0065363.
 XX
 PR 23-APR-1999; 99US-0065363.
 XX
 PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
 XX
 PI Holtzman DA;
 XX
 DR WPI: 2000-013240/01.
 DR N-PSDB: Z37133.
 XX
 PT Novel polynucleotides and polypeptides used to modulate a variety of
 PT cellular processes
 XX
 PS Example 4; Fig 6; 120pp; English.
 XX
 CC Y52138-Y52140 are amino acid sequences of spliced variants of TANGO 125
 CC (T125). The T125 protein (Y52137) has two epidermal growth factor
 CC (EGF)-like domains at amino acids 107-134 and 141-176 and is predicted to
 CC have a molecular weight of approximately 30kD. T125 is predicted to have
 CC no transmembrane domains and appears to be a secreted protein. The
 CC sequences of all variants of T125 are used in the invention to create
 CC antibodies which selectively bind to T125. The T125 polypeptide is used

CC to modulate a variety of cellular processes. It can be used to produce
 CC fusion proteins. The protein may also be used to produce antibodies, and
 CC to identify T125 antagonists and agonists. The T125 polynucleotides, and
 CC polypeptides, homologues and antibodies can be used in screening assays;
 CC disorders. The T125 polynucleotides can be used to express the protein;
 CC to detect T125 mRNA; to detect genetic alterations in the T125 gene; in
 CC forensic biology; and as a source of primers and probes.
 SQ Sequence 247 AA;

Query Match 87.8%; Score 43; DB 21; Length 247;
 Best Local Similarity 87.5%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGCVOPG 8
 11111111
 Db 68 ggcvcvpg 95

RESULT 10
 W88382
 ID W88382 standard; Protein: 254 AA.
 AC W88382;
 XX
 DT 26-APR-1999 (first entry)
 XX
 DE Human neuro-growth factor-like protein Zneul mature polypeptide.
 KW Zneul-1; neuro-growth factor-like protein; human; breast cancer;
 KW glioblastoma; pituitary adenoma; Alzheimer's disease; therapy;
 KW nerve regeneration; haematopoiesis; fertility; contraception;
 KW antibody.
 XX
 OS Homo sapiens.
 XX

Key Location/Qualifiers
 FT 1..85
 FT /note="Hydrophilic domain (HSM1), homologous to
 FT an HSMHC3W5A domain"
 FT Domain 86..116
 FT /note="epidermal growth factor-like domain 1"
 FT Domain 117..158
 FT /note="epidermal growth factor-like domain 2"
 FT Domain 159..254
 FT /note="domain HSM2 homologous to an HSMHC3W5A
 FT domain"
 XX

PN W09857983-A2.
 XX
 XX 23-DEC-1998.
 PD
 XX
 PF 18-JUN-1998; 98WO-US12763.
 XX
 PR 18-JUN-1997; 97US-0878322.
 PR 18-JUN-1997; 97US-0050143.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX

PI Blumberg H, Jelinek LJ, Lehner JM, Sheppard PO;
 PI Whitmore TE;
 DR
 XX
 DR WPI: 1999-095324/08.

PT New mammalian Zneul polypeptides - used to e.g. treat Alzheimer's
 PT disease, cancer and to repopulate blood cells
 XX
 PS Claim 6; Page 48-49; 70pp; English.

CC This polypeptide comprises human Zneul mature polypeptide. Zneul
 CC is a new neuro-growth factor-like protein (see also W88381). Its

CC closest human homologue is HSMHC3W5A, a gene in the HLA class III
 CC region, which is contained in a cosmid which contains Notch 4.
 CC Zneul is also homologous to Notch 4 in its EGF-like domains and may
 CC be involved in EGF receptor pathways. Zneul is widely expressed in
 CC adult tissues, with high expression in heart, placenta, spleen,
 CC testis, thyroid, spinal cord and lymph node. Zneul can be used as
 CC a growth, maintenance, or differentiation factor in the spinal
 CC cord, heart, spleen, testis, thyroid and lymph nodes. It may
 CC play a role in breast cancer, glioblastomas, and pituitary
 CC adenomas. Zneul may also be used to treat Alzheimer's disease,
 CC cancer, to repopulate blood cells after chemotherapy, to stimulate
 CC myofibroblast proliferation, stimulate or inhibit growth factors
 CC made in the placenta, in fertility and contraceptive, or to
 CC regenerate nerves. Claimed Zneul polypeptides (see also W88382-97),
 CC including specific domains of Zneul and epitope-bearing portions of
 CC Zneul, can be used to raise specific antibodies for use e.g. in
 CC diagnostic assays.
 SQ Sequence 254 AA;

Query Match 87.8%; Score 43; DB 20; Length 254;
 Best Local Similarity 87.5%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGCVOPG 8
 11111111
 Db 95 ggcvcvpg 102

RESULT 11
 B42204
 ID B42204 standard; Protein: 265 AA.
 XX

DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF1968 polypeptide sequence SEQ ID NO:1936.
 XX

KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnerability; antipsoriatic; antiparkinsonian; neurotrophic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW hypotensive; dermatological; immunosuppressive; antidiabetic;
 KW antiviral; antibacterial; antifungal; antineoplastic; antiinflammatory;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW thrombosis; contraceptive; antiinflammatory disease; coagulation;
 XX

OS Homo sapiens.
 XX
 PN W0200058473-A2.
 PN
 PD 05-OCT-2000.
 PD
 XX

PF 31-MAR-2000; 2000WO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX

PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 PI
 DR WPI: 2000-602362/57.

DR N-PSDB: C76413.
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11: Page 3102: 5507pp; English.
 XX
 CC C7446 to C77606 encode the proteins given in B40237 to B43397, which
 CC represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vlnnary;
 CC antiparkinsonian; neurotrophic; neuroprotective; osteopathic;
 CC anticonvulsant; antithrombotic; immunosuppressant; immunostimulant;
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatological; immunosuppressive; antinflammatory; antibacterial;
 CC antiviral; antifungal; antipneumatic; antihypoid; and antianaemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, antinflammatory disease: to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive.
 CC
 XX
 SQ Sequence 265 AA:
 Query Match 87.88; Score 43; DB 21; Length 265;
 Best Local Similarly 87.58; Pred. NO. 22;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 GGGCVOPG 8
 Db 114 99acvqps 121
 RESULT 12
 Y41769
 ID Y41769 standard: Protein: 273 AA.
 XX
 AC Y41769;
 XX
 DT 07-DEC-1999 (first entry)
 XX
 DE Human PRO213-1 protein sequence.
 XX
 KM Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
 KM probe; blood coagulation disorder; cancer; cellular adhesion disorder;
 KM secreted protein; transmembrane protein.
 XX
 OS Homo sapiens.
 XX
 PN WO946281-A2.
 PD 16-SEP-1999.
 XX
 PE 08-MAR-1999; 99MO-US05028.
 XX
 PR 10-MAR-1998; 98US-0077450.
 PR 11-MAR-1998; 98US-0077632.
 PR 11-MAR-1998; 98US-0077641.
 PR 11-MAR-1998; 98US-0077649.
 PR 12-MAR-1998; 98US-0077791.
 PR 13-MAR-1998; 98US-0078004.
 PR 17-MAR-1998; 98US-0040220.
 PR 20-MAR-1998; 98US-0078886.
 PR 20-MAR-1998; 98US-0078910.
 PR 20-MAR-1998; 98US-0078936.

PR 20-MAR-1998; 98US-0078939.
 PR 23-MAR-1998; 98US-0079294.
 PR 26-MAR-1998; 98US-0079656.
 PR 27-MAR-1998; 98US-0079663.
 PR 27-MAR-1998; 98US-0079664.
 PR 27-MAR-1998; 98US-0079689.
 PR 27-MAR-1998; 98US-0079728.
 PR 27-MAR-1998; 98US-0079786.
 PR 30-MAR-1998; 98US-0079920.
 PR 30-MAR-1998; 98US-0079923.
 PR 31-MAR-1998; 98US-0080105.
 PR 31-MAR-1998; 98US-0080107.
 PR 31-MAR-1998; 98US-0080165.
 PR 31-MAR-1998; 98US-0080194.
 PR 01-APR-1998; 98US-0080327.
 PR 01-APR-1998; 98US-0080328.
 PR 01-APR-1998; 98US-0080333.
 PR 01-APR-1998; 98US-0080334.
 PR 08-APR-1998; 98US-0081049.
 PR 08-APR-1998; 98US-0081070.
 PR 08-APR-1998; 98US-0081071.
 PR 09-APR-1998; 98US-0081195.
 PR 09-APR-1998; 98US-0081203.
 PR 09-APR-1998; 98US-0081229.
 PR 15-APR-1998; 98US-0081817.
 PR 15-APR-1998; 98US-0081838.
 PR 15-APR-1998; 98US-0081952.
 PR 15-APR-1998; 98US-0081955.
 PR 21-APR-1998; 98US-0082568.
 PR 21-APR-1998; 98US-0082569.
 PR 22-APR-1998; 98US-0082700.
 PR 22-APR-1998; 98US-0082704.
 PR 22-APR-1998; 98US-0082804.
 PR 23-APR-1998; 98US-0082767.
 PR 23-APR-1998; 98US-0082796.
 PR 28-APR-1998; 98US-0083336.
 PR 28-APR-1998; 98US-0083322.
 PR 29-APR-1998; 98US-0083392.
 PR 29-APR-1998; 98US-0083495.
 PR 29-APR-1998; 98US-0083496.
 PR 29-APR-1998; 98US-0083499.
 PR 29-APR-1998; 98US-0083500.
 PR 29-APR-1998; 98US-0083545.
 PR 29-APR-1998; 98US-0083554.
 PR 29-APR-1998; 98US-0083558.
 PR 29-APR-1998; 98US-0083559.
 PR 30-APR-1998; 98US-0083742.
 PR 05-MAY-1998; 98US-0084366.
 PR 06-MAY-1998; 98US-0084414.
 PR 06-MAY-1998; 98US-0084441.
 PR 07-MAY-1998; 98US-0084598.
 PR 07-MAY-1998; 98US-0084600.
 PR 07-MAY-1998; 98US-0084627.
 PR 07-MAY-1998; 98US-0084637.
 PR 07-MAY-1998; 98US-0084639.
 PR 07-MAY-1998; 98US-0084640.
 PR 07-MAY-1998; 98US-0084643.
 PR 13-MAY-1998; 98US-0085323.
 PR 13-MAY-1998; 98US-0085338.
 PR 13-MAY-1998; 98US-0085339.
 PR 15-MAY-1998; 98US-0085573.
 PR 15-MAY-1998; 98US-0085579.
 PR 15-MAY-1998; 98US-0085580.
 PR 15-MAY-1998; 98US-0085582.
 PR 15-MAY-1998; 98US-0085689.
 PR 15-MAY-1998; 98US-0085697.
 PR 15-MAY-1998; 98US-0085700.
 PR 15-MAY-1998; 98US-0085704.
 PR 18-MAY-1998; 98US-0086023.
 PR 22-MAY-1998; 98US-0086392.
 PR 22-MAY-1998; 98US-0086414.
 PR 22-MAY-1998; 98US-0086430.
 PR 22-MAY-1998; 98US-0086486.

PR 28-MAY-1998; 98US-0087098.
PR 28-MAY-1998; 98US-0087106.
PR 28-MAY-1998; 98US-0087208.
PR 30-JUL-1998; 98US-0094651.
PR 11-SEP-1998; 98US-0100038.
XX
PA (GETH) GENENTECH INC.
XX
XX
PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
XX
DR WPI: 1999-551358/46.
DR N-PSDB: 234311.
XX
FT New secreted and transmembrane polypeptides and their polynucleotides,
PT useful for treating blood coagulation disorders, cancers and cellular
XX adhesion disorders -
XX
PS Claim 12: Fig 213: 530pp; English.
XX
CC The present invention describes secreted and transmembrane polypeptides
CC and their polynucleotides. The nucleotide sequences are useful as
CC sources of probes, primers, for chromosome mapping, and for generation
CC of antisense sequences. They can also be used to create transgenic
CC animals. The proteins can be used to treat a variety of diseases and
CC disorders, depending on their function. Diseases that may be treated
CC include blood coagulation disorders, cancers and cellular adhesion
CC disorders. They may also be used to raise antibodies. 233891 to
CC 234338, and 241685 to 241774 represent polynucleotide and polypeptide
CC sequence given in the exemplification of the present invention.
XX
SO Sequence 273 AA;

Query Match 87.8%; Score 43; DB 20; Length 273;
Best Local Similarity 87.5%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GCGCVQPC 8
Db 114 gscvcp9 121

RESULT 13
Y41770
ID Y41770 standard; Protein; 273 AA.
XX
AC Y41770;
XX
DT 07-DEC-1999 (first entry)
XX
DE Human PRO1330 protein sequence.
XX
KW Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
KW probe; blood coagulation disorder; cancer; cellular adhesion disorder;
KW secreted protein; transmembrane protein.
XX
OS Homo sapiens.
XX
PN WO9946281-A2.
XX
PD 16-SEP-1999.
XX
PF 08-MAR-1999; 99WO-US05028.
XX
PR 10-MAR-1998; 98US-0077450.
PR 11-MAR-1998; 98US-0077632.
PR 11-MAR-1998; 98US-0077641.
PR 12-MAR-1998; 98US-0077649.
PR 13-MAR-1998; 98US-0077791.
PR 17-MAR-1998; 98US-0040220.
PR 20-MAR-1998; 98US-0078866.
PR 20-MAR-1998; 98US-0078910.

PR 20-MAR-1998; 98US-0078936.
PR 20-MAR-1998; 98US-0078939.
PR 26-MAR-1998; 98US-0079294.
PR 27-MAR-1998; 98US-0079656.
PR 27-MAR-1998; 98US-0079664.
PR 27-MAR-1998; 98US-0079689.
PR 27-MAR-1998; 98US-0079728.
PR 30-MAR-1998; 98US-0079766.
PR 30-MAR-1998; 98US-0079920.
PR 31-MAR-1998; 98US-0080105.
PR 31-MAR-1998; 98US-0080107.
PR 31-MAR-1998; 98US-0080165.
PR 31-MAR-1998; 98US-0080194.
PR 01-APR-1998; 98US-0080327.
PR 01-APR-1998; 98US-0080328.
PR 01-APR-1998; 98US-0080333.
PR 01-APR-1998; 98US-0080334.
PR 08-APR-1998; 98US-0081049.
PR 08-APR-1998; 98US-0081070.
PR 08-APR-1998; 98US-0081071.
PR 09-APR-1998; 98US-0081195.
PR 09-APR-1998; 98US-0081203.
PR 15-APR-1998; 98US-0081229.
PR 15-APR-1998; 98US-0081817.
PR 15-APR-1998; 98US-0081838.
PR 15-APR-1998; 98US-0081952.
PR 21-APR-1998; 98US-0081925.
PR 21-APR-1998; 98US-0082568.
PR 22-APR-1998; 98US-0082569.
PR 22-APR-1998; 98US-0082700.
PR 22-APR-1998; 98US-0082704.
PR 23-APR-1998; 98US-0082804.
PR 23-APR-1998; 98US-0082867.
PR 27-APR-1998; 98US-0082796.
PR 27-APR-1998; 98US-0083336.
PR 29-APR-1998; 98US-0083322.
PR 29-APR-1998; 98US-0083495.
PR 29-APR-1998; 98US-0083496.
PR 29-APR-1998; 98US-0083499.
PR 29-APR-1998; 98US-0083500.
PR 29-APR-1998; 98US-0083501.
PR 29-APR-1998; 98US-0083554.
PR 29-APR-1998; 98US-0083556.
PR 29-APR-1998; 98US-0083559.
PR 30-APR-1998; 98US-0083742.
PR 05-MAY-1998; 98US-0084366.
PR 06-MAY-1998; 98US-0084414.
PR 07-MAY-1998; 98US-0084441.
PR 07-MAY-1998; 98US-0084598.
PR 07-MAY-1998; 98US-0084600.
PR 07-MAY-1998; 98US-0084627.
PR 07-MAY-1998; 98US-0084637.
PR 07-MAY-1998; 98US-0084639.
PR 07-MAY-1998; 98US-0084640.
PR 13-MAY-1998; 98US-0084643.
PR 13-MAY-1998; 98US-0085328.
PR 15-MAY-1998; 98US-0085339.
PR 15-MAY-1998; 98US-0085373.
PR 15-MAY-1998; 98US-0085579.
PR 15-MAY-1998; 98US-0085580.
PR 15-MAY-1998; 98US-0085582.
PR 15-MAY-1998; 98US-0085589.
PR 15-MAY-1998; 98US-0085689.
PR 15-MAY-1998; 98US-0085697.
PR 15-MAY-1998; 98US-0085700.
PR 18-MAY-1998; 98US-0085704.
PR 22-MAY-1998; 98US-0086023.
PR 22-MAY-1998; 98US-0086392.
PR 22-MAY-1998; 98US-0086414.
PR 22-MAY-1998; 98US-0086430.

PR 22-MAY-1998; 98US-0086486.
PR 28-MAY-1998; 98US-0087098.
PR 28-MAY-1998; 98US-0087106.
PR 28-MAY-1998; 98US-0087208.
PR 30-JUL-1998; 98US-0094651.
PR 11-SEP-1998; 98US-0100038.
XX
PA (GETH) GENENTECH INC.
XX
PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
XX
DR MPI: 1999-551358/46.
DR N-PSDB; 234312.
XX
PT New secreted and transmembrane polypeptides and their polynucleotides,
PT useful for treating blood coagulation disorders, cancers and cellular
PT adhesion disorders -
XX
PS Claim 12; Fig 215; 530pp; English.
XX
CC The present invention describes secreted and transmembrane polypeptides
CC and their polynucleotides. The nucleotide sequences are useful as
CC sources of probes, primers, for chromosome mapping, and for generation
CC of antisense sequences. They can also be used to create transgenic
CC animals. The proteins can be used to treat a variety of diseases and
CC disorders, depending on their function. Diseases that may be treated
CC include blood coagulation disorders, cancers and cellular adhesion
CC disorders. They may also be used to raise antibodies. 233891 to
CC 234338, and 241685 to 241774 represent polynucleotide and polypeptide
CC sequence given in the exemplification of the present invention.
XX
SQ Sequence 273 AA;

Query Match 87.8%; Score 43; DB 20; Length 273;
Best Local Similarity 87.5%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGCVOPG 8
DB 114 99scvqpg 121

RESULT 14
Y41771
ID Y41771 standard; Protein: 273 AA.
XX
AC Y41771;
XX
DT 07-DEC-1999 (first entry)
XX
DE Human PRO1449 protein sequence.
XX
KW Human; PRO: EST: expressed sequence tag; PCR primer; hybridisation;
KW probe; blood coagulation disorder; cancer; cellular adhesion disorder;
KW secreted protein; transmembrane protein.
XX
OS Homo sapiens.
XX
PN WO9946281-A2.
XX
PD 16-SEP-1999.
XX
PF 08-MAR-1999; 99WO-US05028.
XX
PR 10-MAR-1998; 98US-0077450.
PR 11-MAR-1998; 98US-0077632.
PR 11-MAR-1998; 98US-0077641.
PR 11-MAR-1998; 98US-0077649.
PR 12-MAR-1998; 98US-0077791.
PR 13-MAR-1998; 98US-0078004.
PR 17-MAR-1998; 98US-0040220.
PR 20-MAR-1998; 98US-0078866.

PR 20-MAR-1998; 98US-0078910.
PR 20-MAR-1998; 98US-0078936.
PR 20-MAR-1998; 98US-0078939.
PR 25-MAR-1998; 98US-0079294.
PR 26-MAR-1998; 98US-0079656.
PR 27-MAR-1998; 98US-0079663.
PR 27-MAR-1998; 98US-0079664.
PR 27-MAR-1998; 98US-0079668.
PR 27-MAR-1998; 98US-0079728.
PR 27-MAR-1998; 98US-0079786.
PR 30-MAR-1998; 98US-0079920.
PR 31-MAR-1998; 98US-0079923.
PR 31-MAR-1998; 98US-0080105.
PR 31-MAR-1998; 98US-0080107.
PR 31-MAR-1998; 98US-0080165.
PR 31-MAR-1998; 98US-0080194.
PR 01-APR-1998; 98US-0080327.
PR 01-APR-1998; 98US-0080328.
PR 01-APR-1998; 98US-0080333.
PR 01-APR-1998; 98US-0080334.
PR 08-APR-1998; 98US-0081049.
PR 08-APR-1998; 98US-0081070.
PR 08-APR-1998; 98US-0081071.
PR 09-APR-1998; 98US-0081195.
PR 09-APR-1998; 98US-0081203.
PR 09-APR-1998; 98US-0081229.
PR 15-APR-1998; 98US-0081817.
PR 15-APR-1998; 98US-0081838.
PR 15-APR-1998; 98US-0081952.
PR 15-APR-1998; 98US-0081955.
PR 21-APR-1998; 98US-0082568.
PR 21-APR-1998; 98US-0082569.
PR 22-APR-1998; 98US-0082700.
PR 22-APR-1998; 98US-0082704.
PR 22-APR-1998; 98US-0082804.
PR 23-APR-1998; 98US-0082767.
PR 23-APR-1998; 98US-0082796.
PR 27-APR-1998; 98US-0083336.
PR 28-APR-1998; 98US-0083322.
PR 29-APR-1998; 98US-0083392.
PR 29-APR-1998; 98US-0083495.
PR 29-APR-1998; 98US-0083496.
PR 29-APR-1998; 98US-0083499.
PR 29-APR-1998; 98US-0083500.
PR 29-APR-1998; 98US-0083545.
PR 29-APR-1998; 98US-0083554.
PR 29-APR-1998; 98US-0083558.
PR 29-APR-1998; 98US-0083559.
PR 30-APR-1998; 98US-0083742.
PR 05-MAY-1998; 98US-0084366.
PR 06-MAY-1998; 98US-0084411.
PR 06-MAY-1998; 98US-0084441.
PR 07-MAY-1998; 98US-0084598.
PR 07-MAY-1998; 98US-0084600.
PR 07-MAY-1998; 98US-0084627.
PR 07-MAY-1998; 98US-0084637.
PR 07-MAY-1998; 98US-0084639.
PR 07-MAY-1998; 98US-0084640.
PR 07-MAY-1998; 98US-0084643.
PR 13-MAY-1998; 98US-0085323.
PR 13-MAY-1998; 98US-0085338.
PR 13-MAY-1998; 98US-0085339.
PR 15-MAY-1998; 98US-0085373.
PR 15-MAY-1998; 98US-0085379.
PR 15-MAY-1998; 98US-0085580.
PR 15-MAY-1998; 98US-0085582.
PR 15-MAY-1998; 98US-0085587.
PR 15-MAY-1998; 98US-0085697.
PR 15-MAY-1998; 98US-0085700.
PR 15-MAY-1998; 98US-0085704.
PR 18-MAY-1998; 98US-0086023.
PR 22-MAY-1998; 98US-0086392.
PR 22-MAY-1998; 98US-0086414.

PR 22-MAY-1998; 98US-0086430.
 PR 28-MAY-1998; 98US-0086486.
 PR 28-MAY-1998; 98US-0087098.
 PR 28-MAY-1998; 98US-0087106.
 PR 30-JUL-1998; 98US-0087208.
 PR 11-SEP-1998; 98US-0094651.
 PR 11-SEP-1998; 98US-0100038.

PA (GETH) GENENTECH INC.

PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
 DR WPI: 1999-551358/46.
 DR N-PSDB; 234313.

XX New secreted and transmembrane polypeptides and their polynucleotides,
 PT useful for treating blood coagulation disorders, cancers and cellular
 XX adhesion disorders -

PS Claim 12; Fig 217; 530pp; English.

CC The present invention describes secreted and transmembrane polypeptides
 CC and their polynucleotides. The nucleotide sequences are useful as
 CC sources of probes, primers, for chromosome mapping, and for generation
 CC of antisense sequences. They can also be used to create transgenic
 CC animals. The proteins can be used to treat a variety of diseases and
 CC disorders, depending on their function. Diseases that may be treated
 CC include blood coagulation disorders, cancers and cellular adhesion
 CC disorders. They may also be used to raise antibodies. 233891 to
 CC 234338, and Y41685 to Y41774 represent polynucleotide and polypeptide
 CC sequence given in the exemplification of the present invention.

XX Sequence 273 AA;

Query Match 87.8%; Score 43; DB 20; Length 273;
 Best Local Similarity 87.5%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCGVPG 8
 11111111
 Db 114 ggcvcvpg 121

RESULT 15

W88381
 ID W88381 standard; protein; 273 AA.

AC W88381.

XX 26-APR-1999 (first entry)

DE Human neuro-growth factor-like protein zneu1.

XX Zneu-1; neuro-growth factor-like protein; human; breast cancer;
 KW glioblastoma; pituitary adenoma; Alzheimer's disease; therapy;
 KW nerve regeneration; hematopoiesis; fertility; contraception;
 XX antibody.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..19 /note="putative signal peptide"

FT Protein 20..273 /note="mature protein"

FT Domain 20..104 /note="hydrophilic domain (HSM1), homologous to
 an HSMHC3W5A domain"

FT Domain 105..135 /note="epidermal growth factor-like domain 1"

FT Domain 136..177 /note="epidermal growth factor-like domain 2"

FT Domain

178..273

/note="domain HSM2 homologous to an HSMHC3W5A domain"

FT W09857983-A2.

XX 23-DEC-1998.

PF 18-JUN-1998; 98WO-US12763.

PR 18-JUN-1997; 97US-0878322.

PR 18-JUN-1997; 97US-0050143.

PA (ZYMO) ZYMOGENETICS INC.

PI Blumberg H, Jellinek LJ, Lehner JW, Sheppard PO;

DR N-PSDB; V84341.

DR WPI: 1999-095324/08.

XX New mammalian zneu1 polypeptides - used to, e.g. treat Alzheimer's
 XX disease, cancer and to repopulate blood cells

PS Claim 6; Page 47-48; 70pp; English.

CC This polypeptide comprises human zneu1, a new neuro-growth factor-
 CC like protein. Its amino acid sequence was deduced from the
 CC nucleotide sequence (see V84341) of a cDNA clone isolated from a
 CC brain cDNA library. Zneu1's closest human homologue is HSMHC3W5A,
 CC which contains Notch 4. Zneu1 is also homologous to Notch 4 in its
 CC EGF-like domains and may be involved in EGF receptor pathways.
 CC Zneu1 is widely expressed in adult tissues, with high expression in
 CC heart, placenta, spleen, testis, thyroid, spinal cord and lymph
 CC node. Zneu1 polypeptide can be used as a growth, maintenance, or
 CC differentiation factor in the spinal cord, heart, spleen, testis,
 CC thyroid and lymph nodes. It may also play a role in breast cancer,
 CC glioblastomas, and pituitary adenomas. Zneu1 may be used to treat
 CC Alzheimer's disease, cancer, to repopulate blood cells after
 CC chemotherapy, to stimulate myofibroblast proliferation, stimulate
 CC or inhibit growth factors made in the placenta, in fertility and
 CC polypeptides (see also W88382-97), including specific domains of
 CC Zneu1 and epitope-bearing portions of Zneu1, can be used to raise
 CC specific antibodies for use e.g. in diagnostic assays.

XX Sequence 273 AA;

Query Match 87.8%; Score 43; DB 20; Length 273;
 Best Local Similarity 87.5%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCGVPG 8
 11111111
 Db 114 ggcvcvpg 121

Search completed: June 13, 2001, 14:25:44
 Job time: 663 sec

Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:10:49 ; Search time 87.97 Seconds
(without alignments)
6.250 Million cell updates/sec

Title: PCT-US01-05825A-21

Perfect score: 49

Sequence: 1 GGCVCVQP 8

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	87.8	145	2 H69051	heterodisulfide re
2	40	81.6	130	1 KRSH3A	keratin high-sulfu
3	40	81.6	131	1 KRSH3A	keratin high-sulfu
4	40	81.6	131	1 KRGT3M	keratin high-sulfu
5	40	81.6	132	1 KRGT3J	keratin high-sulfu
6	40	81.6	278	2 E65020	ethanolamine utill
7	39	79.6	36	2 D31485	Ig heavy chain V r
8	39	79.6	38	2 S33402	Ig heavy chain V r
9	39	79.6	59	2 A27606	Ig heavy chain V-a
10	39	79.6	59	2 S36381	Ig heavy chain V r
11	39	79.6	60	2 S36382	Ig heavy chain V r
12	39	79.6	82	2 C34964	Ig heavy chain pre
13	39	79.6	94	2 PL0120	Ig heavy chain V-I
14	39	79.6	94	2 D25913	Ig heavy chain V r
15	39	79.6	94	2 I67528	CD33 antigen homol
16	39	79.6	95	2 I67527	CD33 antigen homol
17	39	79.6	96	2 PH0873	Ig heavy chain V r
18	39	79.6	97	2 PH0872	Ig heavy chain V r
19	39	79.6	97	2 S44115	Ig heavy chain V r
20	39	79.6	97	2 S26885	Ig heavy chain V r
21	39	79.6	97	2 S26886	Ig heavy chain V r
22	39	79.6	97	2 S26890	Ig heavy chain V r
23	39	79.6	97	2 S46462	Ig heavy chain V r
24	39	79.6	97	2 S26895	Ig heavy chain V r
25	39	79.6	97	2 S26895	Ig heavy chain V r
26	39	79.6	98	2 S54855	Ig heavy chain V r
27	39	79.6	98	2 PL0121	Ig heavy chain V-I
28	39	79.6	98	2 PH0874	Ig heavy chain V-I
29	39	79.6	98	2 PL0123	Ig heavy chain V-I

30	39	79.6	98	2 S26896	Ig heavy chain V r
31	39	79.6	98	2 S29545	Ig heavy chain V r
32	39	79.6	98	2 S29546	Ig heavy chain V r
33	39	79.6	98	2 S26927	Ig heavy chain V r
34	39	79.6	98	2 S26932	Ig heavy chain V r
35	39	79.6	98	2 S29543	Ig heavy chain V r
36	39	79.6	98	2 S26891	Ig heavy chain V r
37	39	79.6	98	2 S26894	Ig heavy chain V r
38	39	79.6	98	2 S26889	Ig heavy chain V r
39	39	79.6	98	2 S26933	Ig heavy chain V r
40	39	79.6	98	2 S26934	Ig heavy chain V r
41	39	79.6	98	2 S54856	Ig heavy chain V r
42	39	79.6	100	2 PL0122	Ig heavy chain V-I
43	39	79.6	100	2 S69886	Ig heavy chain V r
44	39	79.6	100	2 S26925	Ig heavy chain V r
45	39	79.6	100	2 S26926	Ig heavy chain V r

ALIGNMENTS

RESULT 1
H69051
heterodisulfide reductase Hrd related protein - Methanobacterium thermoautotrophicum
C:Species: Methanobacterium thermoautotrophicum
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: H69051
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.
; Liu, D.; Spadafora, R.; Vicalre, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jilwan, K.I.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
A:Reference number: A69000; MUID:9803514
A:Accession: H69051
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-145 <MTH>
A:Cross-references: GB:AE000803; GB:AE000666; NID:g2621179; PIDN:AMB84645.1; PID:g262
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH139

Query Match 87.8%; Score 43; DB 2; Length 145;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCVCVQP 7
Db 100 GGCVCVQP 106

RESULT 2
KRSH3A
keratin high-sulfur matrix protein IIA3A - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 16-Aug-1996
C:Accession: A02841
R:Swart, L.S.; Haylett, T.
Biochem. J. 133, 641-654, 1973
A:Title: Studies on the high-sulfur proteins of reduced merino wool. Amino acid sequ
A:Reference number: A90269; MUID:74022242
A:Accession: A02841
A:Molecule type: protein
A:Residues: 1-130 <SWA>
A:Experimental source: Merino wool
C:Superfamily: keratin high-sulfur matrix protein IIA
C:Keywords: duplication; hair

Query Match 81.6%; Score 40; DB 1; Length 130;
Best Local Similarity 85.7%; Pred. No. 6.3;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCVOP 7
 Db 15 GGGCLOP 21

RESULT 3

KRSH3

keratin high-sulfur matrix protein IIA3 - sheep

N:Alternate names: M2.6 protein

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 16-Aug-1996

C:Accession: A02840

R:Swart, L.S.; Haylett, T.

Biochem. J. 133, 641-654, 1973

A:Title: Studies on the high-sulphur proteins of reduced merino wool. Amino acid sequenc

A:Reference number: A90269; MUID:74022242

A:Accession: A02840

A:Molecule type: protein

A:Residues: 1-131 <SMA>

A:Experimental source: Merino wool

C:Comment: Wool and hair consist of microfibrils embedded in a rigid matrix of other pro

C:Superfamily: keratin high-sulfur matrix protein IIA

C:Keywords: duplication; hair

Query Match 81.6%; Score 40; DB 1; Length 131;
 Best Local Similarity 85.7%; Pred. No. 6.3;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGCVOP 7
 Db 15 GGGCLOP 21

RESULT 4

KRG3M

keratin high-sulfur matrix protein IIA3, minor component - goat

N:Alternate names: M2.6 protein

C:Species: Capra aegagrus hircus (domestic goat)

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Aug-1996

C:Accession: B92978; A02840

R:Joubert, F.J.

J. S. Afr. Chem. Inst. 28, 250-263, 1975

A:Title: Studies on the high-sulphur proteins of reduced mohair. The isolation and the a

A:Reference number: A92978

A:Accession: B92978

A:Molecule type: protein

A:Residues: 1-131 <TOD>

A:Experimental source: Angora breed

C:Comment: Wool and hair consist of microfibrils embedded in a rigid matrix of other pro

C:Superfamily: keratin high-sulfur matrix protein IIA

C:Keywords: duplication; hair

Query Match 81.6%; Score 40; DB 1; Length 131;
 Best Local Similarity 85.7%; Pred. No. 6.3;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGCVOP 7
 Db 15 GGGCLOP 21

RESULT 5

KRG3J

keratin high-sulfur matrix protein IIA3, major component - goat

N:Alternate names: M2.6 protein

C:Species: Capra aegagrus hircus (domestic goat)

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Aug-1996

C:Accession: A92978; A02840

R:Joubert, F.J.

J. S. Afr. Chem. Inst. 28, 250-263, 1975

A:Title: Studies on the high-sulphur proteins of reduced mohair. The isolation and th
 A:Reference number: A92978
 A:Accession: A92978
 A:Molecule type: protein
 A:Residues: 1-132 <TOD>
 A:Experimental source: Angora breed
 C:Comment: Wool and hair consist of microfibrils embedded in a rigid matrix of other
 C:Superfamily: keratin high-sulfur matrix protein IIA
 C:Keywords: duplication; hair

Query Match 81.6%; Score 40; DB 1; Length 132;
 Best Local Similarity 85.7%; Pred. No. 6.4;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGCVOP 7
 Db 15 GGGCLOP 21

RESULT 6

E65020

ethanolamine utilization protein EutY - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 08-Oct-1999

C:Accession: E65020

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.,

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: E65020

A>Status: Preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-278 <BLAT>

A:Cross-references: GB:AE000332; GB:U00096; NID:91788789; PIDN:AACT5507.1; PID:917887

A:Experimental source: strain K-12, substrain MG1655

C:Genetics: eutJ

Query Match 81.6%; Score 40; DB 2; Length 278;
 Best Local Similarity 75.0%; Pred. No. 12;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGCVOP 8
 Db 229 GGGCQPG 236

RESULT 7

D31485

1g heavy chain V regions (3-24 and 5-14) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 30-May-1997

C:Accession: D31485

R:Bedzyk, W.D.; Johnson, L.S.; Riordan, G.S.; Voss Jr., E.W.

J. Biol. Chem. 264, 1565-1569, 1989

A:Title: Comparison of variable region primary structures within an anti-fluorescein

A:Reference number: A31485; MUID:89109167

A:Accession: D31485

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-36 <ABD>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotrimer; immunoglobulin

Query Match 79.6%; Score 39; DB 2; Length 36;
 Best Local Similarity 87.5%; Pred. No. 3.1;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGCVOP 8

Db 8 GCGLVQPG 15

RESULT 8
S33402

Ig heavy chain V region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 17-Jul-1998

C:Accession: S33402; S36385

R:Kettleborough, C.A.; Saldanha, J.; Ansell, K.H.; Bendig, M.M.

Eur. J. Immunol. 23, 206-211, 1993

A:Title: Optimization of primers for cloning libraries of mouse immunoglobulin genes

A:Reference number: S33391; MUID:93122092

A:Accession: S33402

A:Molecule type: mRNA

A:Residues: 1-38 <KET>

A:Cross-references: EMBL:X73019

A:Experimental source: strain BALB/c

R:Ansell, K.H.

submitted to the EMBL Data Library, April 1993

A:Reference number: S36376

A:Accession: S36385

A:Molecule type: mRNA

A:Residues: 1-30 <ANS>

A:Cross-references: EMBL:X73019

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:1-38/Domain: immunoglobulin homology (fragment) <IMM>

Query Match 79.6%: Score 39; DB 2; Length 38;

Best Local Similarity 87.5%: Pred. No. 3.3;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCGCVQPG 8
||| ||||
Db 6 GCGLVQPG 13

RESULT 9
A27606

Ig heavy chain V-a region (p26.9p2) - rabbit (fragments)

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 08-Mar-1989 #sequence_revision 30-Jun-1991 #text_change 16-Aug-1996

C:Accession: A27606

R:Currier, S.J.; Gallarda, J.L.; Knight, K.L.

J. Immunol. 140, 1651-1659, 1988

A:Title: Partial molecular genetic map of the rabbit V-H chromosomal region.

A:Reference number: A27606; MUID:88154464

A:Accession: A27606

A:Molecule type: DNA

A:Residues: 1-59 <CUR>

A:Cross-references: GB:M19706

A>Note: the authors translated the codon TCC for residue 28 as Asp

C:Genetics:

A:introns: 30/3

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match 79.6%: Score 39; DB 2; Length 59;

Best Local Similarity 87.5%: Pred. No. 4.7;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCGCVQPG 8
||| ||||
Db 8 GCGLVQPG 15

RESULT 10
S36381

Ig heavy chain V region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 23-Jul-1999

C:Accession: S36381; S33396

R:Ansell, K.H.

submitted to the EMBL Data Library, April 1993

A:Reference number: S36376

A:Accession: S36381

A:Molecule type: mRNA

A:Residues: 1-59 <ANS>

A:Cross-references: EMBL:X73012; NID:q295879; PIDN:CAAS1498.1; PID:q939941

R:Kettleborough, C.A.; Saldanha, J.; Ansell, K.H.; Bendig, M.M.

Eur. J. Immunol. 23, 206-211, 1993

A:Title: Optimization of primers for cloning libraries of mouse immunoglobulin genes

A:Reference number: S33391; MUID:93122092

A:Accession: S33396

A:Molecule type: mRNA

A:Residues: 3-43 <KET>

A:Cross-references: EMBL:X73012

A:Experimental source: strain BALB/c

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:14-59/Domain: immunoglobulin homology (fragment) <IMM>

Query Match 79.6%: Score 39; DB 2; Length 59;

Best Local Similarity 87.5%: Pred. No. 4.7;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCGCVQPG 8
||| ||||
Db 10 GCGLVQPG 17

RESULT 11
S36382

Ig heavy chain V region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 23-Jul-1999

C:Accession: S36382; S33397

R:Ansell, K.H.

submitted to the EMBL Data Library, April 1993

A:Reference number: S36376

A:Accession: S36382

A:Molecule type: mRNA

A:Residues: 1-60 <ANS>

A:Cross-references: EMBL:X73011; NID:q295878; PIDN:CAAS1497.1; PID:q939940

R:Kettleborough, C.A.; Saldanha, J.; Ansell, K.H.; Bendig, M.M.

Eur. J. Immunol. 23, 206-211, 1993

A:Title: Optimization of primers for cloning libraries of mouse immunoglobulin genes

A:Reference number: S33391; MUID:93122092

A:Accession: S33397

A:Molecule type: mRNA

A:Residues: 1-33 <KET>

A:Cross-references: EMBL:X73011

A:Experimental source: strain BALB/c

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:8-60/Domain: immunoglobulin homology (fragment) <IMM>

Query Match 79.6%: Score 39; DB 2; Length 60;

Best Local Similarity 87.5%: Pred. No. 4.8;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCGCVQPG 8
||| ||||
Db 1 GCGLVQPG 8

RESULT 12
C34964

Ig heavy chain precursor V-II region (Ab21) - human

C:Species: Homo sapiens (man)

C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-May-1997
 C/Accession: C34964
 R:Sanz, I.; Casali, P.; Thomas, J.W.; Notkins, A.L.; Capra, J.D.
 J. Immunol. 142, 4054-4061, 1989
 A>Title: Nucleotide sequences of eight human natural autoantibody V-H regions reveals af
 A/Reference number: A92830; MUID:89235232
 A/Accession: C34964
 A/Molecule type: mRNA
 A/Residues: 1-82 <SAN>
 A/Cross-references: GB:M26996
 A/Note: the sequence shown differs from the authors' translation after residue 56
 C/Superfamily: immunoglobulin V region; immunoglobulin homology
 C/Keywords: heterotetramer; immunoglobulin

Query Match
 Best Local Similarity 79.6%; Score 39; DB 2; Length 82;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGCVQPG 8
 ||| ||||
 DB 27 GGGVQPG 34

RESULT 13
 PL0120
 Ig heavy chain V-III region (TD-Vo) - human (fragment)
 C/Species: Homo sapiens (man)
 C/Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 16-Aug-1996
 C/Accession: PL0120
 R:Bird, J.; Gallili, N.; Link, M.; Stites, D.; Sklar, J.
 J. Exp. Med. 168, 229-245, 1988
 A>Title: Continuing rearrangement but absence of somatic hypermutation in immunoglobulin
 A/Reference number: PL0116; MUID:88286083
 A/Accession: PL0120
 A/Molecule type: mRNA
 A/Residues: 1-94 <BIR>
 A/Experimental source: B cells from patient TD with acute lymphoblastic leukemia, ALL
 A/Note: the sequence shows the V region (TD-Vo) from a nonproductive DNA rearrangement
 C/Superfamily: immunoglobulin V region; immunoglobulin homology
 C/Keywords: acute lymphoblastic leukemia; heterotetramer; immunoglobulin
 F:11-94/Region: complementarity-determining 1
 F:49-65/Region: complementarity-determining 2

Query Match
 Best Local Similarity 79.6%; Score 39; DB 2; Length 94;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGCVQPG 8
 ||| ||||
 DB 8 GGGVQPG 15

RESULT 14
 D25913
 Ig heavy chain V region (BFL1) - mouse (fragment)
 C/Species: Mus musculus (house mouse)
 C/Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 16-Aug-1996
 C/Accession: D25913
 R:Lawler, A.M.; Lin, P.S.; Gearhart, P.J.
 Proc. Natl. Acad. Sci. U.S.A. 84, 2454-2458, 1987
 A>Title: Adult B-cell repertoire is biased toward two heavy-chain variable-region genes
 A/Reference number: A94148; MUID:87175692
 A/Accession: D25913
 A/Molecule type: DNA
 A/Residues: 1-94 <LAW>
 C/Superfamily: immunoglobulin V region; immunoglobulin homology
 C/Keywords: heterotetramer; immunoglobulin
 F:11-94/Domain: immunoglobulin homology <LAW>

Best Local Similarity 87.5%; Pred. No. 7;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGCVQPG 8
 ||| ||||
 DB 4 GGGVQPG 11

RESULT 15
 I67528
 CD33 antigen homolog - mouse (fragment)
 C/Species: Mus sp. (mouse)
 C/Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jan-2000
 C/Accession: I67528
 R:Chies, J.A.; Lembezat, M.P.; Freitas, A.A.
 Eur. J. Immunol. 24, 1657-1664, 1994
 A>Title: Entry of B lymphocytes into the persistent cell pool in non-immunized mice 1
 A/Reference number: I53392; MUID:94298870
 A/Accession: I67528
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-94 <RES>
 A/Cross-references: GB:S71357; NID:9550040
 C/Genetics:
 A/Gene: Ig VHx24b
 C/Superfamily: immunoglobulin V region; immunoglobulin homology
 F:11-94/Domain: immunoglobulin homology <LAW>

Query Match
 Best Local Similarity 79.6%; Score 39; DB 2; Length 94;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGCVQPG 8
 ||| ||||
 DB 4 GGGVQPG 11

Search completed: June 13, 2001, 14:23:10
 Job time: 741 sec

Query Match 79.6%; Score 39; DB 2; Length 94;

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:21:50 ; Search time 51.57 Seconds

(without alignments)
5.314 Million cell updates/sec

Title: PCT-US01-05825A-21

Perfect score: 49
Sequence: 1 GGGCYQPG 8

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SWISSProt_39:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	40	81.6	130 1	KRA3_SHEEP
2	40	81.6	131 1	KRA3_SHEEP
3	40	81.6	132 1	KRA3_SHEEP
4	40	81.6	278 1	KRA3_SHEEP
5	40	81.6	279 1	KRA3_SHEEP
6	39	79.6	111 1	KRA3_SHEEP
7	39	79.6	113 1	KRA3_SHEEP
8	39	79.6	113 1	KRA3_SHEEP
9	39	79.6	113 1	KRA3_SHEEP
10	39	79.6	113 1	KRA3_SHEEP
11	39	79.6	113 1	KRA3_SHEEP
12	39	79.6	113 1	KRA3_SHEEP
13	39	79.6	113 1	KRA3_SHEEP
14	39	79.6	113 1	KRA3_SHEEP
15	39	79.6	113 1	KRA3_SHEEP
16	39	79.6	113 1	KRA3_SHEEP
17	39	79.6	113 1	KRA3_SHEEP
18	39	79.6	113 1	KRA3_SHEEP
19	39	79.6	113 1	KRA3_SHEEP
20	39	79.6	113 1	KRA3_SHEEP
21	39	79.6	113 1	KRA3_SHEEP
22	39	79.6	113 1	KRA3_SHEEP
23	39	79.6	113 1	KRA3_SHEEP
24	39	79.6	113 1	KRA3_SHEEP
25	39	79.6	113 1	KRA3_SHEEP
26	39	79.6	113 1	KRA3_SHEEP
27	39	79.6	113 1	KRA3_SHEEP
28	39	79.6	113 1	KRA3_SHEEP
29	39	79.6	113 1	KRA3_SHEEP
30	39	79.6	113 1	KRA3_SHEEP
31	39	79.6	113 1	KRA3_SHEEP
32	39	79.6	113 1	KRA3_SHEEP
33	39	79.6	113 1	KRA3_SHEEP

34	39	79.6	122 1	HV20_MOUSE	P01789 mus musculu
35	39	79.6	122 1	HV21_MOUSE	P01790 mus musculu
36	39	79.6	123 1	HV18_MOUSE	P01787 mus musculu
37	39	79.6	123 1	HV19_MOUSE	P01788 mus musculu
38	39	79.6	123 1	HV22_MOUSE	P01791 mus musculu
39	39	79.6	123 1	HV23_MOUSE	P01792 mus musculu
40	39	79.6	123 1	HV24_MOUSE	P01793 mus musculu
41	39	79.6	123 1	HV25_MOUSE	P01794 mus musculu
42	39	79.6	126 1	HV3K_HUMAN	P01772 homo sapien
43	39	79.6	136 1	HV16_MOUSE	P01783 mus musculu
44	39	79.6	142 1	HV01_RAT	P01805 rattus norv
45	39	79.6	144 1	HV26_MOUSE	P01795 mus musculu

ALIGNMENTS

RESULT 1	KRA3_SHEEP	STANDARD:	PRT:	130 AA.
ID	P02443;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	01-AUG-1991 (Rel. 19, Last annotation update)			
DE	KERATIN, HIGH-SULFUR MATRIX PROTEIN, IIA3A.			
OS	Ovis aries (Sheep).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Caprinae; Ovis.			
OX	NCBI_TaxID=9940;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE=7402242; PubMed=4584026;			
RA	Swart L.S., Haylett T.;			
RT	*Studies on the high-sulphur proteins of reduced merino wool. Amino			
RT	acid sequence of protein SCMB-3A3.*;			
RL	Biochem. J. 133:641-654(1973).			
CC	-1- FUNCTION: THE KERATIN PRODUCTS OF MAMMALIAN EPIDERMAL DERIVATIVES			
CC	SUCH AS WOOL AND HAIR CONSIST OF MICROFIBRILS EMBEDDED IN A RIGID			
CC	MATRIX OF OTHER PROTEINS. THE MATRIX PROTEINS INCLUDE THE HIGH-			
CC	SULFUR AND HIGH-TYROSINE KERATINS, HAVING MOLECULAR WEIGHTS OF			
CC	6-20 KDA, WHEREAS THE MICROFIBRILS CONTAIN THE LARGER, LOW-SULFUR			
CC	KERATINS (40-56 KDA).			
CC	-1- MISCELLANEOUS: THE SOURCE OF THIS KERATIN IS MERINO WOOL.			
DR	PIR: A02841; KR3A3.			
DR	InterPro: IPR002494; -			
DR	Pfam: PF01500; keratin_B2; 1.			
KW	keratin.			
SQ	SEQUENCE 130 AA; 13894 MW; 8ACF3719FBB8361 CRC64;			
QY	1 GGGCYQPG 7			
DB	15 GGGCLOP 21			
Query Match	81.6%; Score 40; DB 1; Length 130;			
Best Local Similarity	85.7%; Pred. No. 2.6;			
Matches	6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			
RESULT 2	KRA3_SHEEP	STANDARD:	PRT:	131 AA.
ID	P02441;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	01-AUG-1991 (Rel. 19, Last annotation update)			
DE	KERATIN, HIGH-SULFUR MATRIX PROTEIN, IIA3A.			
OS	Ovis aries (Sheep).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Caprinae; Ovis.			
OX	NCBI_TaxID=9940;			

RP [1]
 RX MEDLINE-74022242; PubMed-4584026;
 RA Swart L.S., Haylett T.;
 RT "Studies on the high-sulphur proteins of reduced Merino wool. Amino
 RL acid sequence of protein SCMB-3A3.";
 CC Blochem. J. 133:641-654(1973).
 CC -1- FUNCTION: THE KERATIN PRODUCTS OF MAMMALIAN EPIDERMAL DERIVATIVES
 CC SUCH AS WOOL AND HAIR CONSIST OF MICROFIBRILS EMBEDDED IN A RIGID
 CC MATRIX OF OTHER PROTEINS. THE MATRIX PROTEINS INCLUDE THE HIGH-
 CC SULFUR AND HIGH-TYROSINE KERATINS, HAVING MOLECULAR WEIGHTS OF
 CC 6-20 KDA, WHEREAS THE MICROFIBRILS CONTAIN THE LARGER, LOW-SULFUR
 CC KERATINS (40-56 KDA).
 CC -1- MISCELLANEOUS: THE SOURCE OF THIS PROTEIN IS MERINO WOOL.
 DR PIR: A02840; KRSH3.
 DR InterPro: IPR002494;
 DR Pfam: PF01500; Keratin_B2.1.
 DR Keratin.
 KM
 SO SEQUENCE 131 AA; 14163 MW; 8C0B56C0070B1C50 CRC64;

Query Match 81.6%; Score 40; DB 1; Length 131;
 Best Local Similarity 85.7%; Pred. No. 2.6;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGCVP 7
 DB 15 GGGCVP 21

RESULT 3
 ID KRA3_CAPHI STANDARD; PRT; 132 AA.
 AC P02442;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DE 15-DEC-1998 (Rel. 37, Last annotation update)
 DE KERATIN, HIGH-SULFUR MATRIX PROTEIN, IIL3 (M2.6 PROTEIN).
 OS Capra hircus (Goat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Caprinae; Capra.
 OX NCBI_TaxID=9925;
 RN [1]
 RP SEQUENCE.
 RC STRAIN-SOUTH AFRICAN ANGORA;
 RA Joubert F.J.;
 RT "Studies on the high-sulphur proteins of reduced mohair. The
 RL isolation and the amino acid sequence of protein SCMB-M2.6.";
 CC J. S. Afr. Chem. Inst. 28:250-263(1975).
 CC -1- FUNCTION: THE KERATIN PRODUCTS OF MAMMALIAN EPIDERMAL DERIVATIVES
 CC SUCH AS WOOL AND HAIR CONSIST OF MICROFIBRILS EMBEDDED IN A RIGID
 CC MATRIX OF OTHER PROTEINS. THE MATRIX PROTEINS INCLUDE THE HIGH-
 CC SULFUR AND HIGH-TYROSINE KERATINS, HAVING MOLECULAR WEIGHTS OF
 CC 6-20 KDA, WHEREAS THE MICROFIBRILS CONTAIN THE LARGER, LOW-SULFUR
 CC KERATINS (40-56 KDA).
 CC PIR: A02978; KRGT3J.
 DR PIR: B92978; KRGT3M.
 DR InterPro: IPR002494;
 DR Pfam: PF01500; Keratin_B2.1.
 DR Keratin.
 KM
 SO SEQUENCE 132 AA; 14255 MW; 6C0C08BFA13AB28B CRC64;

Query Match 81.6%; Score 40; DB 1; Length 132;
 Best Local Similarity 85.7%; Pred. No. 2.6;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGCVP 7
 DB 15 GGGCVP 21

RESULT 4
 ID EUTJ_ECOLI STANDARD; PRT; 278 AA.
 AC P77277;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE ETHANOLAMINE UTILIZATION PROTEIN EUTJ.
 GN EUTJ.
 OS Escherichia coli.
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-97426617; PubMed-9278503;
 RA Blatter F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RT Mau B., Shao Y.;
 RL "The complete genome sequence of Escherichia coli K-12.";
 RN Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RA Alba H., Baba T., Fujita K., Hayashi K., Honjo A., Horluchi T.,
 RA Ikemoto K., Inada T., Isono K., Itoh T., Kanai K., Kasai H.,
 RA Kashimoto K., Kim S., Kimura S., Kitagawa M., Kitakawa M., Makino K.,
 RA Masuda S., Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y.,
 RA Nishimoto H., Nishio Y., Oshima T., Saito N., Samped G., Seki Y.,
 RA Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- PATHWAY: ETHANOLAMINE UTILIZATION.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC or send an email to license@isb-sib.ch).

DR EMBL: AE000332; AAC75507.1;
 DR EMBL: D90874; BAA16332.1;
 DR EcoGene; EG14184; eutJ
 SO SEQUENCE 278 AA; 30082 MW; 8F92B55DE54ED9FF CRC64;

Query Match 81.6%; Score 40; DB 1; Length 278;
 Best Local Similarity 75.0%; Pred. No. 4.8;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGCVP 8
 DB 229 GGGCVP 236

RESULT 5
 ID EUTJ_SALTY STANDARD; PRT; 279 AA.
 AC P41794;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE ETHANOLAMINE UTILIZATION PROTEIN EUTJ.
 GN EUTJ.
 OS Salmonella typhimurium.
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Salmonella.
 OX NCBI_TaxID=602;


```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 140285;
RX MEDLINE-9517314; PubMed=7868611;
RA Stojilkovic I., Baemler A.J., Heffron F.;
RT "Ethanolamine utilization in Salmonella typhimurium: nucleotide
sequence, protein expression, and mutational analysis of the ccha
cchB gene cluster."
RL J. Bacteriol. 177:1357-1366(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-LT2;
RX MEDLINE-99395039; PubMed=10464203;
RA Kofoid E.C., Rappleye C.A., Stojilkovic I., Roth J.R.;
RT "The 17-gene ethanolamine (eut) operon of Salmonella typhimurium
encodes five homologues of carboxysome shell proteins."
RL J. Bacteriol. 181:5317-5329(1999).
CC -1- PANTHAY: ETHANOLAMINE UTILIZATION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U18560; AAA80210.1; -.
CC DR EMBL: AF093749; AAC78119.1; -.
CC DR Sycenga; SG10554; eutJ.
SO SEQUENCE 279 AA; 30018 MW; 28B8CD89141D8D90 CRC64;

Query Match
Best Local Similarity 81.6%; Score 40; DB 1; Length 279;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGCVQPG 8
DB 230 GGSCMQPG 237

RESULT 6
HV35_MOUSE STANDARD; PRT: 111 AA.
AC P01804;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-II REGION HPC76 (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-81013937; PubMed=6251474;
RA Bernard O., Gough N.M.;
RT "Nucleotide sequence of immunoglobulin heavy chain joining segments
between translocated VH and mu constant regions genes."
RL Proc. Natl. Acad. Sci. U.S.A. 77:3630-3634(1980).
CC -1- MISCELLANEOUS: THE SEQUENCE OF THE FIRST 197 RESIDUES OF THE C
REGION WAS ALSO DETERMINED & DIFFERS IN ONLY 3 POSITIONS FROM THE
CC CORRESPONDING PORTION OF THE MOUSE MOPC 104E MU CHAIN.
CC PIR: A02074; MHMS76.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 1g; 1.
KW Immunoglobulin V region.
FT NON_TER 1
FT SEQUENCE 111 AA; 12304 MW; OEDE98EC7348056A CRC64;
SO SEQUENCE 111 AA; 12304 MW; OEDE98EC7348056A CRC64;

```

```

Query Match
Best Local Similarity 79.6%; Score 39; DB 1; Length 113;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGCVQPG 8
DB 8 GGGLVQPG 10

RESULT 7
HV27_MOUSE STANDARD; PRT: 113 AA.
AC P01796;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-II REGION A4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE-78158406; PubMed=417344;
RA Vrana M., Rudikoff S., Potter M.;
RT "Sequence variation among heavy chains from inulin-binding myeloma
proteins."
RL Proc. Natl. Acad. Sci. U.S.A. 75:1957-1961(1978).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS INULIN.
CC PIR: A93818; AVMSAB.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 1g; 1.
KW Immunoglobulin V region.
FT DISULFID 22
FT NON_TER 113
FT SEQUENCE 113 AA; 12675 MW; 76658C121C598285 CRC64;

Query Match
Best Local Similarity 79.6%; Score 39; DB 1; Length 113;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGCVQPG 8
DB 8 GGGLVQPG 15

RESULT 8
HV28_MOUSE STANDARD; PRT: 113 AA.
AC P01797;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-II REGION U61.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE-78158406; PubMed=417344;
RA Vrana M., Rudikoff S., Potter M.;
RT "Sequence variation among heavy chains from inulin-binding myeloma
proteins."
RL Proc. Natl. Acad. Sci. U.S.A. 75:1957-1961(1978).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS INULIN.
CC PIR: B93818; AVMS61.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 1g; 1.
KW Immunoglobulin V region.

```

FT DISULFID 22 98 BY SIMILARITY.
NON_TER 113 113
SEQUENCE 113 AA; 12671 MW; 76634C121C598285 CRC64;

Query Match
Best Local Similarity 79.6%; Score 39; DB 1; Length 113;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGCVQPG 8
DB 8 GGGVQPG 15

RESULT 9

HV29_MOUSE STANDARD; PRT: 113 AA.

ID HV29_MOUSE
AC P01798;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1993 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION E109.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090.
RN [1]
RP SEQUENCE.

RX MEDLINE=78158406; PubMed=417344;
RA Vrana M., Rudikoff S., Potter M.;

RT "Sequence variation among heavy chains from inulin-binding myeloma
proteins.";

RL Proc. Natl. Acad. Sci. U.S.A. 75:1957-1961(1978).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
BINDS INULIN.

DR PIR: C93818; AVMS09.
DR InterPro: IPR003006; -

DR Pfam: PF00047; 15; 1.
KW Immunoglobulin V region.

FT DISULFID 22 98 BY SIMILARITY.
NON_TER 113 113

SEQUENCE 113 AA; 12647 MW; E550F2F30EDB129B CRC64;

Query Match

Best Local Similarity 79.6%; Score 39; DB 1; Length 113;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGCVQPG 8
DB 8 GGGVQPG 15

RESULT 10

HV30_MOUSE STANDARD; PRT: 113 AA.

ID HV30_MOUSE
AC P01799;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1993 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION ABE-47N.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090.
RN [1]
RP SEQUENCE.

RX MEDLINE=7134726; PubMed=402936;
RA Vrana M., Rudikoff S., Potter M.;

RT "Heavy-chain variable-region sequence from an inulin-binding myeloma
protein.";

RL Biochemistry 16:1170-1175(1977).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT

CC BINDS INULIN.

DR PIR: A90400; AVMS87.
DR InterPro: IPR003006; -

DR Pfam: PF00047; 15; 1.
KW Immunoglobulin V region.

FT DISULFID 22 98 BY SIMILARITY.
NON_TER 113 113

SEQUENCE 113 AA; 12675 MW; 76658C16C779845E CRC64;

Query Match
Best Local Similarity 79.6%; Score 39; DB 1; Length 113;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGCVQPG 8
DB 8 GGGVQPG 15

RESULT 11

HV31_MOUSE STANDARD; PRT: 113 AA.

ID HV31_MOUSE
AC P01800;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-OCT-1993 (Rel. 27, Last sequence update)
DE 15-JUL-1993 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION 1957.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090.
RN [1]
RP SEQUENCE.

RX MEDLINE=81216632; PubMed=6787122;
RA Rudikoff S., Potter M.;

RT "Immunoglobulin heavy chains from anti-inulin myeloma proteins:
evidence for a new heavy chain joining segment.";

RL J. Immunol. 127:191-194(1981).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
BINDS INULIN.

DR PIR: A92810; AVMS87.
DR InterPro: IPR003006; -

DR Pfam: PF00047; 15; 1.
KW Immunoglobulin V region.

FT DISULFID 22 98 BY SIMILARITY.
NON_TER 113 113

SEQUENCE 113 AA; 12732 MW; 26618F626B59859E CRC64;

Query Match

Best Local Similarity 79.6%; Score 39; DB 1; Length 113;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGCVQPG 8
DB 8 GGGVQPG 15

RESULT 12

HV34_MOUSE STANDARD; PRT: 113 AA.

ID HV34_MOUSE
AC P01803;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1993 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V REGION AMPC1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090.
RN [1]
RP SEQUENCE.

RX MEDLINE=81216632; PubMed=6787122;.

RA Rudkoff S., Potter M.;
 RT "Immunoglobulin heavy chains from anti-inulin myeloma proteins:
 evidence for a new heavy chain joining segment."
 RL J. Immunol. 127:191-194(1981).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
 BINDS INULIN.
 DR PIR: A02073; HVMSAM.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region.
 FT DISULFID 22 98
 FT NON_TER 113 113
 SO SEQUENCE 113 AA; 12691 MW; 7A6D906AA966E9E CRC64;

Query Match 79.6%; Score 39; DB 1; Length 113;
 Best Local Similarity 87.5%; Pred. No. 3.4;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGCVOPG 8
 DB 8 GGGVOPG 15

RESULT 13
 HV32_MOUSE STANDARD; PRT; 115 AA.
 AC P01801;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION J606.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=8209361; PubMed=6798111;
 RA Johnson N., Stankard J., Paul L., Hood L.;
 RT "The complete V domain amino acid sequences of two myeloma inulin-
 binding proteins."
 RT J. Immunol. 128:302-307(1982).
 RL -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
 BINDS INULIN.
 CC PIR: C92811; AVMS06.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region.
 FT DISULFID 22 98
 FT NON_TER 115 115
 SO SEQUENCE 115 AA; 12810 MW; B67AD6638A121A5F CRC64;

Query Match 79.6%; Score 39; DB 1; Length 115;
 Best Local Similarity 87.5%; Pred. No. 3.4;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGCVOPG 8
 DB 8 GGGVOPG 15

RESULT 14
 HV33_MOUSE STANDARD; PRT; 115 AA.
 AC P01802;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION W3082.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=8209361; PubMed=6798111;
 RA Johnson N., Stankard J., Paul L., Hood L.;
 RT "The complete V domain amino acid sequences of two myeloma inulin-
 binding proteins."
 RT J. Immunol. 128:302-307(1982).
 RL -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
 BINDS INULIN.
 CC PIR: D92811; AVMS82.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region.
 FT DISULFID 22 98
 FT NON_TER 115 115
 SO SEQUENCE 115 AA; 12887 MW; 9B4517648C121C5A CRC64;

Query Match 79.6%; Score 39; DB 1; Length 115;
 Best Local Similarity 87.5%; Pred. No. 3.4;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGCVOPG 8
 DB 8 GGGVOPG 15

RESULT 15
 HV3D_HUMAN STANDARD; PRT; 115 AA.
 AC P01765;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION T1L.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=78005528; PubMed=409716;
 RA Wang A.-C., Wang I.Y., Fudenberg H.H.;
 RT "Immunoglobulin structure and genetics. Identity between variable
 regions of a mu and a gamma2 chain."
 RT J. Biol. Chem. 252:7192-7199(1977).
 RL -1- MISCELLANEOUS: THE SEQUENCES OF THE V REGIONS OF THE HEAVY CHAINS
 OF IGM AND IGG2 ISOLATED FROM A SINGLE PATIENT WITH BICLONAL
 GAMMOPATHY ARE IDENTICAL. THEIR LIGHT CHAINS ARE APPARENTLY ALSO
 IDENTICAL.
 CC PIR: A02048; H3HUTL.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region.
 FT NON_TER 115 115
 SO SEQUENCE 115 AA; 12356 MW; 4DCC67D179F62326 CRC64;

Query Match 79.6%; Score 39; DB 1; Length 115;
 Best Local Similarity 87.5%; Pred. No. 3.4;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGCVOPG 8
 DB 8 GGGVOPG 15

Search completed: June 13, 2001, 14:30:35
 Job time: 525 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:20:39 ; Search time 150.43 Seconds
(without alignments)
6.233 Million cell updates/sec

Title: PCT-US01-05825A-21

Perfect score: 49
Sequence: 1 GGGCVQPG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_15: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_unclassified: *
13: sp_vertebrate: *
14: sp_virus: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	87.8	145	1	026242
2	43	87.8	273	4	Q9UHF1
3	43	87.8	622	5	Q9VHM5
4	39	79.6	15	11	Q9QV16
5	39	79.6	95	4	Q9ULB6
6	39	79.6	113	4	Q9ULB0
7	39	79.6	116	4	Q9ULB3
8	39	79.6	118	4	Q9ULB1
9	39	79.6	121	4	Q9ULB2
10	39	79.6	122	4	Q9ULB4
11	39	79.6	147	4	Q9V509
12	39	79.6	616	2	Q49182
13	39	79.6	1690	5	Q9VXU0
14	39	79.6	1708	5	016039
15	39	79.6	278	11	Q9QXT5
16	38	77.6	534	8	006418
17	38	77.6	542	5	023446
18	38	77.6	704	5	Q9XTM4
19	38	77.6			Q9XTM4 drosophila

20	38	77.6	870	6	002660	002660 bos taurus
21	37	75.5	377	5	Q9U6R4	Q9U6R4 trypanosoma
22	37	75.5	645	3	094106	094106 pneumocysti
23	37	75.5	647	3	094104	094104 pneumocysti
24	36.5	74.5	270	5	P90606	P90606 trypanosoma
25	36	73.5	168	2	Q9S2V7	Q9S2V7 streptomyces
26	36	73.5	417	4	Q9UER3	Q9UER3 homo sapien
27	36	73.5	550	4	Q94883	Q94883 homo sapien
28	36	73.5	583	10	Q22889	Q22889 arabidopsis
29	36	73.5	587	11	Q88831	Q88831 rattus norv
30	36	73.5	588	4	Q9Y5N2	Q9Y5N2 homo sapien
31	36	73.5	1193	5	Q19617	Q19617 caenorhabdi
32	36	73.5	1235	5	Q9VFE3	Q9VFE3 drosophila
33	36	73.5	1689	4	P78528	P78528 homo sapien
34	36	73.5	2233	5	Q94711	Q94711 paramecium
35	36	73.5	2395	5	Q27167	Q27167 paramecium
36	36	73.5	2717	5	Q94710	Q94710 paramecium
37	36	73.5	2721	5	Q76973	Q76973 paramecium
38	36	73.5	4097	4	Q13327	Q13327 homo sapien
39	36	73.5	4128	4	P78527	P78527 homo sapien
40	35	71.4	57	4	Q9UP29	Q9UP29 homo sapien
41	35	71.4	131	4	Q9ULB8	Q9ULB8 homo sapien
42	35	71.4	206	13	Q90802	Q90802 gallus gall
43	35	71.4	223	2	Q25533	Q25533 helicobacte
44	35	71.4	223	2	Q92KY6	Q92KY6 helicobacte
45	35	71.4	245	5	Q46338	Q46338 naegleria s

ALIGNMENTS

RESULT 1
ID 026242 PRELIMINARY: PRT: 145 AA.
AC 026242;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE HETERODISULFIDE REDUCTASE HDRD RELATED PROTEIN.
GN MTHL39.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanobacterium.
OX NCBI_TaxID=2166;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DELTA H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzaden R., Blakey D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lunn W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jivani N., Caruso A., Bush D., Saefer H., Patwell D., Prabhakar S.,
RA McQuigley S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RL deltaH: functional analysis and comparative genomics.";
DR EMBL: AE000803; AAB84645.1;--
DR INTERPRO: IPR000194;--
DR PROSITE: PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
SO SEQUENCE 145 AA; 16448 MW; 694CDACE1A6560A8 CRC64;

Query Match 87.8% Score 43; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGCVQVP 7
|||||
Db 100 GGGCVQVP 106

```

RESULT 2
Q9UHF1 PRELIMINARY: PRT: 273 AA.
AC Q9UHF1:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE NCCH-LIKE PROTEIN.
GN ZN01.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON NCBI_Taxid=9606;
RP SEQUENCE FROM N.A.
RA Sheppard P., Jellinek L., Whitmore T., Blumberg H., Lehner J.,
RA O'Hara P.,
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF186111; AAF01429.1;
DR HSSP; P35555; 1EMN.
DR INTERPRO; IPR000152;
DR INTERPRO; IPR000561;
DR INTERPRO; IPR001881;
DR PFAM; PF00008; EGF_2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS01187; EGF_CA; 1.
SQ SEQUENCE 273 AA; 29617 MW; 5740B845ED5A988 CRC64;

Query Match
Best Local Similarity 87.8%; Score 43; DB 4; Length 273;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGCVCVOPG 8
Db 114 GGSCVOPG 121

RESULT 3
Q9VMH5 PRELIMINARY: PRT: 622 AA.
AC Q9VMH5:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE KR-H1 PROTEIN.
GN KR-H1.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_Taxid=7227;
RP SEQUENCE FROM N.A.
RA STRAIN-BERKELEY.
RA MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sultion G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Peiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Adayani A., An H.-J., Andrews-Pinnkott C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Burits K.C., Busan D.A., Butler H., Brokstein P., Brotlier P.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

```

```

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Tatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Shue B.C., Siden-Kiamos I., Saunders R.D.C., Scheeler F., Shen H.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhou X., Zhu S., Zhu X.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou Q., Zheng L.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
DR EMBL; AEO03612; AAF52343.1;
DR HSSP; P08047; 1SP2.
DR FLYBASE; FBgn0028420; Kr-h1.
DR INTERPRO; IPR000822;
DR PFAM; PF00096; zef-C2H2; 6.
DR PROSITE; PS00028; zef-C2H2; 6.
SQ SEQUENCE 622 AA; 68241 MW; 01E2B29ED64B9EB8 CRC64;

Query Match
Best Local Similarity 87.8%; Score 43; DB 5; Length 622;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCVCVOPG 8
Db 554 GGSCLEPG 561

RESULT 4
Q9QV16 PRELIMINARY: PRT: 15 AA.
AC Q9QV16:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE PROLACTIN-BINDING PROTEIN (FRAGMENT).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_Taxid=10118;
RP SEQUENCE.
RA MEDLINE-95094032; PubMed-8000909;
RA Cohen H., Cohen O., Gagnon J.;
RT "Serum prolactin-binding protein (PRL-BP) of human and rat are
RT identified as IgG."
RL C. R. Acad. Sci., III, Sci. Vie 317:293-298(1994).
DR HSSP; P01789; IMCP.
SQ SEQUENCE 15 AA; 1469 MW; 35ED2512FF3FA369 CRC64;

Query Match
Best Local Similarity 79.6%; Score 39; DB 11; Length 15;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGCVCVOPG 8
Db 8 GGGLVOPG 15

```

RESULT 5
O9ULB6 PRELIMINARY: PRT: 95 AA.
AC O9ULB6:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE IMNOCLOBULIN HEAVY CHAIN (FRAGMENT).
GN
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tange Y., Kayano H.;
RT "Human VH gene sequence."
RL Submitted (NCV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB035268; BAA87067.1; -.
DR HSSP: P01772; 2F84.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; 19; 1.
FT NON_TER 1 1
FT NON_TER 95 95
SO SEQUENCE 95 AA; 10527 MW; 90A8C6D16D2574A CRC64;

Query Match 79.6%; Score 39; DB 4; Length 95;
Best Local Similarity 87.5%; Pred. No. 7.3;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 GGGCVQPG 8
DB 7 GGGVQPG 14

RESULT 6
O9UL90 PRELIMINARY: PRT: 113 AA.
AC O9UL90:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035024; AAD56260.1; -.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; 19; 1.
FT NON_TER 1 1
FT NON_TER 113 113
SO SEQUENCE 113 AA; 12437 MW; ED57FDD19086D07F CRC64;

Query Match 79.6%; Score 39; DB 4; Length 113;
Best Local Similarity 87.5%; Pred. No. 8.6;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 GGGCVQPG 8
DB 8 GGGVQPG 15

RESULT 7
O9UL93 PRELIMINARY: PRT: 116 AA.
AC O9UL93:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035021; AAD56257.1; -.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; 19; 1.
FT NON_TER 1 1
FT NON_TER 116 116
SO SEQUENCE 116 AA; 12434 MW; ODA0348154DD6061 CRC64;

Query Match 79.6%; Score 39; DB 4; Length 116;
Best Local Similarity 87.5%; Pred. No. 8.8;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 GGGCVQPG 8
DB 7 GGGVQPG 14

RESULT 8
O9UL91 PRELIMINARY: PRT: 118 AA.
AC O9UL91:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035023; AAD56259.1; -.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; 19; 1.
FT NON_TER 1 1
FT NON_TER 118 118
SO SEQUENCE 118 AA; 12843 MW; D0633949F2AC149D CRC64;

Query Match 79.6%; Score 39; DB 4; Length 118;
Best Local Similarity 87.5%; Pred. No. 8.9;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 GGGCVQPG 8
DB 8 GGGVQPG 15

```

RESULT 9
090L72 PRELIMINARY; PRT; 118 AA.
AC 090L72;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT)
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalls N.N., Berney S.M.,
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR HSSP; P01772; 2FB4.
DR INTERPRO: IPR003006;
DR PFAM; PF00047; 19; 1.
FT NON_TER 1
FT SEQUENCE 118 AA; 12872 MW; B4DIA5944B2DSCCA CRC64;

```

```

Query Match
Best Local Similarity 79.6%; Score 39; DB 4; Length 118;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 GGGCVQPG 8
Db 8 GGGVQPG 15

```

```

RESULT 10
090L71 PRELIMINARY; PRT; 121 AA.
AC 090L71;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT)
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalls N.N., Berney S.M.,
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR HSSP; P01772; 2FB4.
DR INTERPRO: IPR003006;
DR PFAM; PF00047; 19; 1.
FT NON_TER 1
FT SEQUENCE 121 AA; 13154 MW; 2F045CCPA5D50736 CRC64;

```

```

Query Match
Best Local Similarity 79.6%; Score 39; DB 4; Length 121;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCVQPG 8

```

```

Db 8 GGGVQPG 15
RESULT 11
090L84 PRELIMINARY; PRT; 122 AA.
AC 090L84;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT)
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalls N.N., Berney S.M.,
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR HSSP; P01772; 2FB4.
DR INTERPRO: IPR003006;
DR PFAM; PF00047; 19; 1.
FT NON_TER 1
FT SEQUENCE 122 AA; 13579 MW; 36054DA136545B8 CRC64;

```

```

Query Match
Best Local Similarity 79.6%; Score 39; DB 4; Length 122;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 GGGCVQPG 8
Db 8 GGGVQPG 15

```

```

RESULT 12
09Y509 PRELIMINARY; PRT; 147 AA.
AC 09Y509;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE VH3 PROTEIN (FRAGMENT)
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96071149; PubMed=7475288;
RA Cao J., Vescho R.A., Kettig M.B., Hong C.H., Kim A., Lee J.C.,
RT "A CD10-positive subset of malignant cells is identified in multiple
RL Leukemia 9:1948-1953(1995).
DR EMBL; S80860; AA014339.1;
DR HSSP; P01772; 2FB4.
DR INTERPRO: IPR003006;
DR PFAM; PF00047; 19; 1.
FT NON_TER 1
FT SEQUENCE 147 AA; 15768 MW; 8489FCANA7BC925C CRC64;

```

```

Query Match
Best Local Similarity 79.6%; Score 39; DB 4; Length 147;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```



```

Matches      7;  Conservative      0;  Mismatches      1;  Indels      0;  Gaps      0;

OY      1  GCGCVQPG 8
      ||| ||||
Db      8  GCGVQPG 15

RESULT 13
O49182      PRELIMINARY;      PRT;      616 AA.
AC O49182;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE DNA SEQUENCE, ORFs 1 AND 2.
OS Mycobacterium fortitulum.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_Taxid=1766;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92311251; Pubmed=1615063;
RA Labidi A., Mardis E., Roe B.A., Wallace R.J. Jr.;
RT "Cloning and DNA sequence of the Mycobacterium fortitulum var fortitulum
RT plasmid PAL5000."
RL Plasmid 37:130-140(1992).
DR EMBL: M60875; ANA25373.1; -.
DR INTERPRO: IPR000345; -.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
SQ SEQUENCE      616 AA;      66784 MW;      715EBE768D9A856 CRC64;

```

```

Query Match      79.6%; Score 39; DB 2; Length 616;
Best Local Similarity      85.7%; Pred. No. 42;
Matches      6;  Conservative      1;  Mismatches      0;  Indels      0;  Gaps      0;

OY      1  GCGCVQPG 7
      ||||| |
Db      335 GCGCVRP 341

```

```

RESULT 14
O9VXUO      PRELIMINARY;      PRT;      1690 AA.
AC O9VXUO;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE AC13E PROTEIN.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RX MEDLINE=20196006; Pubmed=10731132;
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Arlitt J.E., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Bayendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burris K.C., Busam D.A., Butler H., Cadenot L.B., Davies P.,
RA Chertis J.M., Cawley S., Dahlke C., Davenport E., Center A., Chandra I.,
RA de Palacios B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

```

```

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson K.A., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003499; AAF48468.1; -.
DR HSSP: P26769; IAB8.
DR FLYBASE: FBgn0022710; AC13E.
DR INTERPRO: IPR001054; -.
DR PFAM: PF00211; guanylate_cyc; 2.
SQ SEQUENCE      1690 AA;      186868 MW;      B762050CFA10DDF0 CRC64;

```

```

Query Match      79.6%; Score 39; DB 5; Length 1690;
Best Local Similarity      75.0%; Pred. No. 11e+02;
Matches      6;  Conservative      0;  Mismatches      2;  Indels      0;  Gaps      0;

OY      1  GCGCVQPG 8
      |||| |
Db      798 GGGCCSPG 805

```

```

RESULT 15
O16039      PRELIMINARY;      PRT;      1708 AA.
AC O16039;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ADENYLYL CYCLASE ISOFORM DAC9.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97431620; Pubmed=9287125;
RA Iourgenko V., Kilot B., Cann M.J., Levin L.R.;
RT "Cloning and characterization of a Drosophila adenylyl cyclase
RT homologous to mammalian type IX."
RL FEBS Lett. 413:104-108(1997)
DR EMBL: AF005630; AAB70469.1; -.
DR HSSP: P26769; IAB8.
DR FLYBASE: FBgn0022710; AC35C.
DR INTERPRO: IPR001054; -.
DR PFAM: PF00211; guanylate_cyc; 2.
SQ SEQUENCE      1708 AA;      189139 MW;      AC1A450CF856D5BFB CRC64;

```

```

Query Match      79.6%; Score 39; DB 5; Length 1708;
Best Local Similarity      75.0%; Pred. No. 1.1e+02;
Matches      6;  Conservative      0;  Mismatches      2;  Indels      0;  Gaps      0;

```

Wed Jun 13 15:00:16 2001

pct-us01-05825a-21.rspt

OY .1 GGGCVQPG 8
|||
Db 798 GGGCCSPG 805

Search completed: June 13, 2001, 14:29:40
Job time: 541 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:16:39 ; Search time 78.71 Seconds
(Without alignments)
1.953 Million cell updates/sec

Title: PCT-US01-05825A-21

Perfect score: 49

Sequence: 1 GGCGVQPG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	79.6	15	1 US-08-331-398A-66	Sequence 66, Appl
2	39	79.6	15	2 US-08-652-558-42	Sequence 42, Appl
3	39	79.6	15	2 US-08-331-397B-66	Sequence 66, Appl
4	39	79.6	15	2 US-08-759-804A-65	Sequence 65, Appl
5	39	79.6	18	1 US-08-331-398A-55	Sequence 55, Appl
6	39	79.6	18	1 US-08-401-908-14	Sequence 14, Appl
7	39	79.6	18	2 US-08-331-397B-55	Sequence 55, Appl
8	39	79.6	18	2 US-08-759-804A-54	Sequence 54, Appl
9	39	79.6	20	2 US-08-859-931A-2	Sequence 2, Appl
10	39	79.6	20	2 US-08-053-451B-113	Sequence 113, App
11	39	79.6	26	1 US-08-471-780C-80	Sequence 80, Appl
12	39	79.6	26	1 US-08-467-282B-80	Sequence 80, Appl
13	39	79.6	26	2 US-08-471-282A-80	Sequence 80, Appl
14	39	79.6	26	2 US-08-466-710C-80	Sequence 80, Appl
15	39	79.6	26	3 US-08-468-739C-80	Sequence 80, Appl
16	39	79.6	29	1 US-08-471-780C-119	Sequence 119, App
17	39	79.6	29	1 US-08-467-282B-119	Sequence 119, App
18	39	79.6	29	2 US-08-471-282A-119	Sequence 119, App
19	39	79.6	29	2 US-08-466-710C-119	Sequence 119, App
20	39	79.6	29	3 US-08-468-739C-119	Sequence 119, App
21	39	79.6	30	1 US-07-988-925-7	Sequence 75, Appl
22	39	79.6	30	1 US-07-977-696C-75	Sequence 75, Appl
23	39	79.6	30	1 US-08-129-830B-75	Sequence 75, Appl
24	39	79.6	30	2 US-08-318-157B-30	Sequence 30, Appl
25	39	79.6	30	2 US-08-318-157B-31	Sequence 31, Appl
26	39	79.6	30	2 US-08-362-780-7	Sequence 7, Appl
27	39	79.6	30	2 US-08-765-783A-92	Sequence 92, Appl

28	39	79.6	30	2 US-08-470-139-17	Sequence 17, Appl
29	39	79.6	67	1 US-08-162-102C-36	Sequence 36, Appl
30	39	79.6	86	2 US-08-053-451B-126	Sequence 126, App
31	39	79.6	87	1 US-08-497-312-21	Sequence 21, Appl
32	39	79.6	87	2 US-08-765-783A-105	Sequence 105, App
33	39	79.6	98	1 US-08-211-202-116	Sequence 116, App
34	39	79.6	98	1 US-08-211-202-118	Sequence 118, App
35	39	79.6	98	1 US-07-942-245-37	Sequence 37, Appl
36	39	79.6	98	2 US-08-428-197-48	Sequence 48, Appl
37	39	79.6	98	2 US-08-665-202-31	Sequence 31, Appl
38	39	79.6	98	5 PCT-US93-10555-48	Sequence 48, Appl
39	39	79.6	108	2 US-08-428-197-4	Sequence 4, Appl
40	39	79.6	108	5 PCT-US93-10555-4	Sequence 4, Appl
41	39	79.6	109	1 US-07-942-245-17	Sequence 17, Appl
42	39	79.6	109	1 US-07-942-245-21	Sequence 21, Appl
43	39	79.6	109	2 US-08-428-197-3	Sequence 3, Appl
44	39	79.6	109	5 PCT-US93-10555-3	Sequence 3, Appl
45	39	79.6	110	1 US-08-211-202-117	Sequence 117, App

ALIGNMENTS

RESULT 1
US-08-331-398A-66
Sequence 66, Application US/08331398A
Patent No. 5608039
GENERAL INFORMATION:
APPLICANT: Pastan, Ira
APPLICANT: Willingham, Mark
APPLICANT: Fitzgerald, David
APPLICANT: Brinkmann, Ulrich
APPLICANT: Pal, Lee
TITLE OF INVENTION: Single Chain B3 Antibody Fusion Proteins
TITLE OF INVENTION: and Their Uses (as amended)
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Stewart Street Plaza
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105-1492
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,398A
FILING DATE: 28-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/767,331
FILING DATE: 30-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/596,289
FILING DATE: 12-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 015280-1261100S
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-5043
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-331-398A-66

Query Match 79.6%; Score 39; DB 1; Length 15;
Best Local Similarity 87.5%; Pred. No. 3.9;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCVOPG 8
Db 8 GGGLVOPG 15

RESULT 2
US-08-652-558-42
Sequence 42, Application US/08652558
Patent No. 5861155
GENERAL INFORMATION:
APPLICANT: LIN, AUGUSTINE YEE-THARN
TITLE OF INVENTION: HUMANIZED ANTIBODIES AND USES
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER & WITCOFF
STREET: 75 STATE STREET, 23RD FLOOR
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,558
FILING DATE: JUNE 6, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB94/00387
FILING DATE: NOVEMBER 21, 1994
ATTORNEY/AGENT INFORMATION:
NAME: YANKRICH, LEON R.
REGISTRATION NUMBER: 30,237
REFERENCE/DOCKET NUMBER: 95,497-L
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-652-558-42

Query Match 79.6%; Score 39; DB 2; Length 15;
Best Local Similarity 87.5%; Pred. No. 3.9;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCVOPG 8
Db 8 GGGLVOPG 15
RESULT 3
US-08-331-397B-66
Sequence 66, Application US/08331397B
Patent No. 5981726
GENERAL INFORMATION:
APPLICANT: Pastan, Ira
APPLICANT: Benhar, Itai
TITLE OF INVENTION: Chimeric and Mutationally Stabilized Tumor-Specific Antibody Fragments, Fusion Proteins, and Uses

TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Stewart Street Plaza
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105-1492
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,397B
FILING DATE: 28-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/767,331
FILING DATE: 30-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/596,289
FILING DATE: 12-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 015280-1261200S
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-331-397B-66

Query Match 79.6%; Score 39; DB 2; Length 15;
Best Local Similarity 87.5%; Pred. No. 3.9;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCVOPG 8
Db 8 GGGLVOPG 15

RESULT 4
US-08-759-804A-65
Sequence 65, Application US/08759804A
Patent No. 5990296
GENERAL INFORMATION:
APPLICANT: Pastan, Ira
APPLICANT: Willingham, Mark
APPLICANT: Fitzgerald, David J.
APPLICANT: Brinkmann, Ulrich
APPLICANT: Pal, Lee
TITLE OF INVENTION: Tumor-Specific Antibody Fragments, Fusion Proteins, and Uses Thereof
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/759,804A
FILING DATE: 03-DEC-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/331,398
FILING DATE: 28-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/767,331
FILING DATE: 30-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/596,289
FILING DATE: 12-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen L.
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 015280-126140US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-759-804A-65

Query Match 79.6%; Score 39; DB 2; Length 15;
Best Local Similarity 87.5%; Pred. No. 3.9;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCVQPG 8
||| ||||
DB 8 GGGLVQPG 15

RESULT 5
US-08-331-398A-55
Sequence 55, Application US/08331398A
Patent No. 5608039
GENERAL INFORMATION:
APPLICANT: Pastan, Ira
APPLICANT: Williamson, Mark
APPLICANT: Fitzgerald, David
APPLICANT: Brinkmann, Ulrich
APPLICANT: Pal, Lee
TITLE OF INVENTION: Single Chain B3 Antibody Fusion Proteins
TITLE OF INVENTION: and Their Uses (as amended)
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Stewart Street Plaza
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105-1492
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,398A
FILING DATE: 28-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/767,331
FILING DATE: 30-SEP-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/596,289
FILING DATE: 12-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 015280-126110US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-331-398A-55

Query Match 79.6%; Score 39; DB 1; Length 18;
Best Local Similarity 87.5%; Pred. No. 4.7;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCVQPG 8
||| ||||
DB 8 GGGLVQPG 15

RESULT 6
US-08-401-908-14
Sequence 14, Application US/08401908
Patent No. 5684146
GENERAL INFORMATION:
APPLICANT: Yoshinobu OKUNO et al.
TITLE OF INVENTION: A DNA WHICH CODES FOR THE VARIABLE REGION OF
TITLE OF INVENTION: ANTIBODY
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/401,908
FILING DATE: March 10, 1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 18
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE: heavy chain of C179

US-08-401-908-14

Query Match 79.6%; Score 39; DB 1; Length 18;
Best Local Similarity 87.5%; Pred. No. 4.7;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCVQPG 8
DB 8 GGGVQPG 15

RESULT 7

US-08-331-397B-55
Sequence 55, Application US/08331397B
Patent No. 5981726

GENERAL INFORMATION:

APPLICANT: Pastan, Ira
TITLE OF INVENTION: Chimeric and Mutationally Stabilized Tumor-
TITLE OF INVENTION: Specific Antibody Fragments, Fusion Proteins, and Uses
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew
CITY: San Francisco
STATE: California
COUNTRY: USA

ZIP: 94105-1492

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,397B

FILING DATE: 28-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: US 07/767,331
FILING DATE: 30-SEP-1991

APPLICATION DATA:
APPLICATION NUMBER: US 07/596,289
FILING DATE: 12-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom

REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 015280-126120US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043

INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid

STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: Peptide

US-08-331-397B-55

Query Match 79.6%; Score 39; DB 2; Length 18;
Best Local Similarity 87.5%; Pred. No. 4.7;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCVQPG 8
DB 8 GGGVQPG 15

RESULT 8

US-08-759-804A-54

Sequence 54, Application US/08759804A
Patent No. 5990296

GENERAL INFORMATION:

APPLICANT: Pastan, Ira
APPLICANT: Willingham, Mark
APPLICANT: Fitzgerald, David J.
APPLICANT: Brinkmann, Ulrich
APPLICANT: Pal, Lee

TITLE OF INVENTION: Tumor-Specific Antibody Fragments,
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
CITY: San Francisco
STATE: California
COUNTRY: USA

ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/759,804A
FILING DATE: 03-DEC-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/331,398
FILING DATE: 28-OCT-1994

APPLICATION DATA:
APPLICATION NUMBER: US 07/767,331
FILING DATE: 30-SEP-1991
APPLICATION NUMBER: US 07/596,289
FILING DATE: 12-OCT-1990

ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen L.
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 015280-126140US

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 54:

SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

MOLECULE TYPE: Peptide
US-08-759-804A-54

Query Match 79.6%; Score 39; DB 2; Length 18;
Best Local Similarity 87.5%; Pred. No. 4.7;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCVQPG 8
DB 8 GGGVQPG 15

RESULT 9

US-08-859-931A-2
Sequence 2, Application US/08859931A
Patent No. 5945510

GENERAL INFORMATION:

APPLICANT: PASANO, Alessio
TITLE OF INVENTION: SUBSTANTIALLY PURE ZONULIN, A
TITLE OF INVENTION: PHYSIOLOGICAL MODULATOR OF
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:

ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W., Suite 800
CITY: Washington, D.C.
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,931A
FILING DATE: 21 May 1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: KIT, Gordon
REGISTRATION NUMBER: 30,764
REFERENCE/DOCKET NUMBER: A-6901
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
US-08-859-931A-2

Query Match 79.6%; Score 39; DB 2; Length 20;
Best Local Similarity 87.5%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCVQPG 8
||| ||||
DB 8 GGGIVQPG 15

RESULT 10
US-08-053-451B-113
Sequence 113, Application US/08053451B
Patent No. 5955584
GENERAL INFORMATION:
APPLICANT: Chen, Francis W.
APPLICANT: Dilew, Charles C.
SOFTWARE: Patentin Release #1.0, Version #1.25
TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC
NUMBER OF SEQUENCES: 176
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/053,451B
FILING DATE: 26-Apr-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P.
REGISTRATION NUMBER: 25,227
REFERENCE/DOCKET NUMBER: 7606-033-999
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-854-3660
TELEFAX: 415-854-3694
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: modified-site
LOCATION: 2-3
OTHER INFORMATION: /note- "where xaa = k or q = Lys or
OTHER INFORMATION: Glu"
US-08-053-451B-113

Query Match 79.6%; Score 39; DB 2; Length 20;
Best Local Similarity 87.5%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCVQPG 8
||| ||||
DB 7 GGGIVQPG 14

RESULT 11
US-08-471-780C-80
Sequence 80, Application US/08471780C
Patent No. 5759808
GENERAL INFORMATION:
APPLICANT: Gasterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins devoid of light chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finegan, Henderson, Farbow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,780C
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potler, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Camelus dromedarius
US-08-471-780C-80

Query Match
Best Local Similarity 79.6%; Score 39; DB 1; Length 26;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGCVCVPG 8
11111111
DB 8 GGGLVOPG 15

RESULT 12
US-08-467-282B-80
Sequence 80, Application US/08467282B
Patent No. 5800988
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flinagan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,282B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Camelus dromedarius
US-08-467-282B-80

Query Match
Best Local Similarity 79.6%; Score 39; DB 1; Length 26;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGCVCVPG 8
11111111

DB 8 GGGLVOPG 15

RESULT 13
US-08-471-282A-80
Sequence 80, Application US/08471282A
Patent No. 5840853
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flinagan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,282A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Camelus dromedarius
US-08-471-282A-80

Query Match
Best Local Similarity 79.6%; Score 39; DB 2; Length 26;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGCVCVPG 8
11111111
DB 8 GGGLVOPG 15

RESULT 14
US-08-466-710C-80
Sequence 80, Application US/08466710C
Patent No. 5874341
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett & Dunner
;; STREET: 1300 I Street, N.W.
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20005-3315
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/466,710C
;; FILING DATE:
;; CLASSIFICATION:
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/106,944
;; FILING DATE: 17-AUG-1993
;; APPLICATION NUMBER: FR 92402326.0
;; FILING DATE: 21-AUG-1992
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: FR 93401310.3
;; FILING DATE: 21-MAY-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Potter, Jane E.R.
;; REGISTRATION NUMBER: 33,332
;; REFERENCE/DOCKET NUMBER: 04958.0008-00000
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-408-4000
;; TELEFAX: 202-408-4400
;;
;; INFORMATION FOR SEQ ID NO: 80:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 26 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; ORIGINAL SOURCE:
;; ORGANISM: Camelus dromedarius
;;
;; US-08-466-710C-80
;;
Query Match 79.6%; Score 39; DB 2; Length 26;
Best Local Similarity 87.5%; Pred. No. 6.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 GGGCVQPG 8
||| ||||
DB 8 GGGLVQPG 15
;;
RESULT 15
US-08-468-739C-80
;; Sequence 80, Application US/08468739C
;; Patent No. 6015695
;; GENERAL INFORMATION:
;; APPLICANT: Casterman, Cecile
;; APPLICANT: Hamers, Raymond
;; TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
;; NUMBER OF SEQUENCES: 130
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett & Dunner
;; STREET: 1300 I Street, N.W.
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20005-3315
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;;

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/468,739C
;; FILING DATE: 06-JUN-1995
;; CLASSIFICATION: 530
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/106,944
;; FILING DATE: 17-AUG-1993
;; APPLICATION NUMBER: FR 92402326.0
;; FILING DATE: 21-AUG-1992
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: FR 93401310.3
;; FILING DATE: 21-MAY-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Potter, Jane E.R.
;; REGISTRATION NUMBER: 33,332
;; REFERENCE/DOCKET NUMBER: 04958.0008-00000
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-408-4000
;; TELEFAX: 202-408-4400
;;
;; INFORMATION FOR SEQ ID NO: 80:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 26 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; ORIGINAL SOURCE:
;; ORGANISM: Camelus dromedarius
;;
;; US-08-468-739C-80
;;

Query Match 79.6%; Score 39; DB 3; Length 26;
Best Local Similarity 87.5%; Pred. No. 6.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 GGGCVQPG 8
||| ||||
DB 8 GGGLVQPG 15

Search completed: June 13, 2001, 14:27:06
Job time: 627 sec

Wed Jun 13 15:00:13 2001

pct-us01-05825a-21.ra1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:25:44 ; Search time 150.28 Seconds
(without alignments)
3.043 Million cell updates/sec

Title: PCT-US01-05825A-22

Perfect score: 48

Sequence: 1 GGCCVDDG 8

Scoring table: BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq_0401.*

- 1: /SIDS6/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SIDS6/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SIDS6/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SIDS6/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SIDS6/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SIDS6/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SIDS6/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SIDS6/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SIDS6/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SIDS6/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SIDS6/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SIDS6/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SIDS6/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SIDS6/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SIDS6/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SIDS6/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SIDS6/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SIDS6/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SIDS6/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SIDS6/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SIDS6/gcgdata/geneseq/geneseq/AA2000.DAT.*
- 22: /SIDS6/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	48	100.0	8 21	Peptide antagonist
2	41	85.4	8 21	Peptide antagonist
3	40	83.3	8 21	Peptide antagonist
4	39	81.2	8 21	Peptide antagonist
5	39	81.2	196 20	Amino acid sequenc
6	38	79.2	8 21	Peptide antagonist
7	38	79.2	8 21	Peptide antagonist
8	37	77.1	586 20	Protein encoded by
9	36	75.0	474 21	Arabidopsis thalia
10	36	75.0	493 10	Carbamate hydrolas
11	36	75.0	541 21	Arabidopsis thalia

12	36	75.0	583	21	G31663	Arabidopsis thalia
13	36	75.0	1487	19	W61562	Human type II coll
14	35	72.9	31	20	W88384	Human zneul EGF-11
15	35	72.9	73	20	W88389	Human zneul partia
16	35	72.5	115	21	B41718	Human ORFX ORF1482
17	35	72.5	153	21	B41638	Human ORFX ORF1402
18	35	72.9	158	20	W88388	Human zneul partia
19	35	72.9	169	20	W88390	Human zneul partia
20	35	72.9	224	20	Y59870	Human normal uteru
21	35	72.9	247	21	Y52139	Human TANGO 125b (
22	35	72.9	254	20	W88382	Human neuro-growth
23	35	72.9	265	21	B42204	Human ORFX ORF1968
24	35	72.9	273	20	Y41769	Human PRO213-1 pro
25	35	72.9	273	20	Y41770	Human PRO1330 prot
26	35	72.9	273	20	Y41771	Human PRO1449 prot
27	35	72.9	273	21	W88381	Human neuro-growth
28	35	72.9	273	21	B44325	Human PRO213-1 pro
29	35	72.9	273	21	B44326	Human PRO1330 prot
30	35	72.9	273	21	B44327	Human PRO1449 prot
31	35	72.9	273	21	B18673	Amino acid sequenc
32	35	72.9	273	21	B18674	Amino acid sequenc
33	35	72.9	273	21	B18675	Amino acid sequenc
34	35	72.9	273	21	B24042	Human PRO213 prote
35	35	72.9	273	21	B24043	Human PRO1330 prot
36	35	72.9	273	21	B24044	Human PRO1449 prot
37	35	72.9	273	21	B01376	Neuron-associated
38	35	72.9	273	21	Y52137	Human TANGO 125 (T
39	35	72.9	295	21	Y41685	Human PRO213 prote
40	35	72.9	295	21	B44241	Human PRO213 (UNQ1
41	35	72.9	307	21	B41644	Human ORFX ORF1408
42	34	70.8	599	17	W16597	Corn acetohydroxy
43	34	70.8	599	17	W16598	Corn acetohydroxy
44	34	70.8	599	17	W16599	Corn acetohydroxy
45	34	70.8	599	17	W03692	Corn acetohydroxy

ALIGNMENTS

RESULT 1	
ID	Y79126 standard; Peptide; 8 AA.
Y79126	
AC	Y79126;
XX	
DT	05-JUN-2000 (first entry)
XX	
DE	Peptide antagonist of zonulin.
XX	
KW	Zonulin; antagonist; zonula occludens toxin receptor;
KW	blood-brain barrier; antiinflammatory; cerebroprotective;
KW	neuroprotective; dermatological; anticulcer; antiviral;
KW	antibacterial; cytoskeletal; anti-HIV; vulnery; antiallergic;
KW	hypotensive; immunosuppressive; antiparasitic; vasotropic;
KW	gastrointestinal inflammation; therapy.
XX	
OS	Synthetic.
XX	
PN	WO200007609-A1.
XX	
PD	17-FEB-2000.
XX	
PF	28-JUL-1999; 99WO-US16683.
XX	
PR	03-AUG-1998; 98US-0127815.
XX	
PA	(UYMA-) UNIV MARYLAND BALTIMORE.
XX	
PI	Fasano A;
XX	
DR	WPI; 2000-205565/18.
XX	
PT	New peptide antagonist of zonulin useful as antiinflammatory agent for

PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 PS Claim 1; Page 48; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.

Sequence 8 AA:

Query Match 100.0%; Score 48; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGCYVDG 8
 DB 1 99gcvcqg 8

RESULT 2
 Y79125
 ID Y79125 standard; Peptide; 8 AA.
 AC Y79125;
 XX

DT 05-JUN-2000 (first entry)
 XX

DE Peptide antagonist of zonulin.
 XX

KW Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; anticulcer; antiviral;
 KW antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.
 KM

XX Synthetic.
 OS

XX WO200007609-A1.
 PN

XX 17-FEB-2000.
 PD

XX 28-JUL-1999; 99WO-US16683.
 PF

XX 03-AUG-1998; 98US-0127815.
 PR

XX (UYMA-) UNIV MARYLAND BALTIMORE.
 PA
 XX

PI Fasano A;
 XX WPI; 2000-205565/18.
 DR

PT New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 PS Claim 1; Page 47; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.

Sequence 8 AA:

Query Match 85.4%; Score 41; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGCYVDG 8
 DB 1 99gcvcqg 8

RESULT 3
 Y79114
 ID Y79114 standard; Peptide; 8 AA.
 AC Y79114;
 XX

DT 05-JUN-2000 (first entry)
 XX

DE Peptide antagonist of zonulin.
 XX

KW Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; anticulcer; antiviral;
 KW antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.
 KM

XX Synthetic.
 OS

XX WO200007609-A1.
 PN

XX 17-FEB-2000.
 PD

XX 28-JUL-1999; 99WO-US16683.
 PF

XX 03-AUG-1998; 98US-0127815.
 XX (UYMA-) UNIV MARYLAND BALTIMORE.
 XX Fasano A;
 XX WPI: 2000-205565/18.
 DR New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -
 XX
 PS Claim 1: Page 44; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
 CC (Z), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.

SO Sequence 8 AA;
 Query Match 83.3%; Score 40; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGCVODG 8
 1 |||||
 Db 1 ggcvcvqdg 8

RESULT 4
 Y79118
 ID Y79118 standard; Peptide: 8 AA.
 XX Y79118;
 AC
 XX 05-JUN-2000 (first entry)
 DT
 XX Peptide antagonist of zonulin.
 DE
 XX zonulin; antagonist; zonula occludens toxin receptor;
 KM blood-brain barrier; antiinflammatory; cerebroprotective;
 KM neuroprotective; dermatological; antitumor; antiviral;
 KM antibacterial; cytoskeletal; anti-HIV; vulnereary; antiallergic;
 KM hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KM gastrointestinal inflammation; therapy.
 XX
 OS Synthetic.

PN WO200007609-A1.
 XX 17-FEB-2000.
 PD
 XX 28-JUL-1999; 99WO-US16683.
 XX
 PF 03-AUG-1998; 98US-0127815.
 XX
 PR (UYMA-) UNIV MARYLAND BALTIMORE.
 XX Fasano A;
 XX WPI: 2000-205565/18.
 DR New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -
 XX
 PS Claim 1: Page 45; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
 CC (Z), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.

SO Sequence 8 AA;
 Query Match 81.2%; Score 39; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGCVODG 8
 1 |||||
 Db 1 ggcvcvqdg 8

RESULT 5
 Y36831
 ID Y36831 standard; Protein: 196 AA.
 XX Y36831;
 AC
 XX 07-OCT-1999 (first entry)
 DT
 XX Amino acid sequence of a Chlamydia trachomatis protein.
 DE
 XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
 KM paratrachoma; inclusion conjunctivitis; genital disease; perhepatitis;
 KM nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
 KM bartolinitis; pneumopathy; venereal lymphogranulomatosis.

```

XX OS Chlamydia trachomatis.
XX PN WO9928475-A2.
XX PD 10-JUN-1999.
XX PF 27-NOV-1998; 98WO-IB01939.
XX PR 04-NOV-1998; 98US-0107077.
XX PR 28-NOV-1997; 97FR-0015041.
XX PR 17-DEC-1997; 97FR-0016034.
XX PA (GERT ) GENSET.
XX PI Griffiths R;
XX DR WPI; 1999-371125/31.
XX PT Genome sequence of Chlamydia trachomatis
XX PS Disclosure: Page 710-711; 1755pp; English.
XX CC Y36754-Y37949 are encoded by open reading frames (ORFs) of the genome of
XX CC Chlamydia trachomatis (see 201425). The polypeptides can be used as
XX CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
XX CC can also be used to control growth of the microorganism. Chlamydia
XX CC trachomatis is responsible for a large number of diseases, e.g. eye
XX CC diseases such as conventional trachoma, nongonococcal trachoma,
XX CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
XX CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
XX CC peritonitis, bartonellosis; pneumonia in breast feeding infants;
XX CC and venereal lymphogranulomatosis. The polypeptides of the invention
XX CC may be of use in treating these diseases.
XX SQ Sequence 196 AA;

Query Match
Best Local Similarity 81.2%; Score 39; DB 20; Length 196;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGCVQDG 8
DB 150 ggycvqdg 157

RESULT 6
ID Y79122 standard; Peptide: 8 AA.
XX AC Y79122;
XX DT 05-JUN-2000 (first entry)
XX DE Peptide antagonist of zonulin.
XX KW Zonulin: antagonist; zonula occludens toxin receptor;
XX KW blood-brain barrier; antiinflammatory; cerebroprotective;
XX KW neuroprotective; dermatological; antiulcer; antiviral;
XX KW antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
XX KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
XX KW gastrointestinal inflammation; therapy.
XX OS Synthetic.
XX PN WO200007609-A1.
XX PD 17-FEB-2000.
XX PF 28-JUL-1999; 99WO-US16683.
XX PR 03-AUG-1998; 98US-0127815.

```

```

XX (UYMA-) UNIV MARYLAND BALTIMORE.
XX FASANO A;
XX WPI; 2000-205565/18.
XX DR
XX PT New peptide antagonist of zonulin useful as antiinflammatory agent for
XX PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
XX PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis.
XX PS Claim 1; Page 46; 69pp; English.
XX
XX This present sequence is that of a peptide antagonist of zonulin
XX CC (2), one of 25 such peptides (see Y79105-29) of the invention,
XX CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
XX CC physiologically modulate the opening of mammalian tight junctions
XX CC (TJ). The peptide antagonists are based on a common motif of ZOT
XX CC binding and human zonulins, which is believed to be critical for receptor
XX CC and human zonulins, which is believed to be critical for receptor
XX CC binding. They can be prepared by chemical synthesis or by use of
XX CC recombinant DNA techniques. The peptide antagonists are used as an
XX CC antiinflammatory agents in the treatment of gastrointestinal
XX CC inflammation, where they bind to the ZOT receptor in the intestine
XX CC and yet does not physiologically modulate the opening of TJ in the
XX CC intestine. Gastrointestinal inflammation conditions give rise to
XX CC increased intestinal permeability and the peptide is useful for
XX CC treating intestinal conditions that cause protein losing enteropathy
XX CC caused by infection, e.g. Clostridium difficile infection,
XX CC enterocolitis, shigellosis, viral gastroenteritis, parasite
XX CC infestation, bacterial overgrowth, whipple's disease, diseases with
XX CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
XX CC collagenous colitis, inflammatory bowel disease, diseases marked by
XX CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
XX CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
XX CC correction of congenital heart disease with Fontan's operation,
XX CC mucosal diseases without ulceration, e.g. Menetrier's disease,
XX CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
XX CC e.g. systemic lupus erythematosus or food allergies, primarily to
XX CC milk.
XX SQ Sequence 8 AA;

```

```

Query Match
Best Local Similarity 79.2%; Score 38; DB 21; Length 8;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGCVQDG 8
DB 1 gglcvqdg 8

RESULT 7
ID Y79128 standard; Peptide: 8 AA.
XX AC Y79128;
XX DT 05-JUN-2000 (first entry)
XX DE Peptide antagonist of zonulin.
XX KW Zonulin: antagonist; zonula occludens toxin receptor;
XX KW blood-brain barrier; antiinflammatory; cerebroprotective;
XX KW neuroprotective; dermatological; antiulcer; antiviral;
XX KW antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
XX KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
XX KW gastrointestinal inflammation; therapy.
XX OS Synthetic.
XX PN WO200007609-A1.

```

PD 17-FEB-2000.
 XX 28-JUL-1999; 99WO-US16683.
 XX 03-AUG-1998; 98US-0127815.
 XX (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PA Fasano A;
 PI
 XX WPI: 2000-205565/18.
 DR
 XX New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 PT
 XX Claim 1; Page 48; 69pp: English.
 PS
 XX This present sequence is that of a peptide antagonist of zonulin
 CC (2). One of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infection, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 CC
 XX Sequence 8 AA:
 SQ
 Query Match 79.2%; Score 38; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. NO. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGCVCVDC 8
 Db 1 99JLVQDQ 8
 Y23621
 ID Y23621 standard; Protein: 586 AA.
 XX Y23621;
 AC
 XX
 DT 06-SEP-1999 (first entry)
 XX
 DE Protein encoded by a reading frame of the porcine circovirus genome.
 XX
 KW MAP: piglet fatal wasting disease; vaccine; circovirus infection;
 KW gene therapy.
 XX Porcine circovirus.
 OS
 XX
 FH Key Location/Qualifiers

FT Misc-difference 1..586
 FT /note= "all x residues are encoded by stop codons"
 XX
 XX FR2772047-A1.
 XX
 XX 11-JUN-1999.
 PD
 XX
 XX 05-DEC-1997; 97FR-0015396.
 PF
 XX
 XX 05-DEC-1997; 97FR-0015396.
 PR
 XX
 XX (NAVE-) CENT NAT ETUD VETERINAIRES & ALIMENTAIRE.
 PA
 XX
 PI Albina E, Arnauld C, Blanchard P, Hutet E, Jestin A;
 PI Le Cam P;
 DR WPI: 1999-360000/31.
 DR N-PSDB: X85593.
 XX
 PT Nucleotide sequence of porcine circovirus MAP - useful in vaccines
 PT against MAP circovirus infection and in gene therapy
 PT
 XX Claim 7; Fig 3; 89pp: French.
 PS
 XX The present sequence is encoded by a reading frame of the positive
 CC strand of a porcine circovirus genome which is associated with MAP.
 CC MAP is the french acronym for piglet fatal wasting disease. The
 CC polypeptides can be used to detect anti-MAP antibodies. The antibodies
 CC can be used to detect MAP antigens. The nucleotide sequences can be used
 CC as probes or primers for detecting MAP nucleic acids. The nucleotide
 CC sequences, polypeptides, vectors, (pseudo)viral particles, transformed
 CC cells and compounds selected by the screening assay can be used in
 CC pharmaceutical compositions. The polypeptides, nucleotide sequences,
 CC vectors and transformed cells can be used in vaccines against MAP
 CC circovirus infection. The vectors, (pseudo)viral particles and
 CC transformed cells can be used for gene therapy.
 CC
 XX Sequence 586 AA:
 SQ
 Query Match 77.1%; Score 37; DB 20; Length 586;
 Best Local Similarity 85.7%; Pred. NO. 3.1e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 GGCVCVDC 8
 Db 552 99CFCQDQ 558
 G31665
 ID G31665 standard; Protein: 474 AA.
 XX G31665;
 AC
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 38065.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 OS
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 XX 25-FEB-2000; 2000EP-0301439.
 PF
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129685.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131444.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136192.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 14-JUN-1999; 99US-0138540.
PR 16-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139465.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141827.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147208.
PR 04-AUG-1999; 99US-0147208.
PR 05-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147483.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151338.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.


```
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 75.0%; Score 36; DB 21; Length 474;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCVCVQ 6
IIIIII
DB 456 ggcvcvq 461

```
RESULT 10
P93413
ID P93413 standard; protein: 493 AA.
AC P93413;
XX
XX 27-APR-1990 (first entry)
DT
XX
DE Carbamate hydrolase.
XX
XX Carbamate hydrolase; Arthrobacter oxidans; phenmedipham;
KW methyl 3-hydroxyphenylcarbamate.
XX
XX Arthrobacter oxidans DSM 4044.
OS
XX
XX EP343100-A.
PN
XX
XX 23-NOV-1989.
PD
XX
XX 17-MAY-1989; 89EP-0730123.
PE
XX
XX 19-MAY-1988; 88DE-3381738.
PR
XX
XX (SCHD ) SCHERING AG.
PA
XX
XX Pohlenz HD, Boldol W;
PI
XX
XX WPI; 1989-341858/47.
DR
XX
XX N-PSDB; N92585.
PT Pure carbamate hydrolase isolation from Arthrobacter oxidans - able to
```

```
PT destroy herbicide phenmedipham, and DNA encoding it, for imparting
PT resistance to plants.
XX
XX Disclosure; Fig 7; 17pp; German.
XX
XX Purified carbamate hydrolase can be used to isolate/identify A. oxidans
CC carbamate hydrolase gene system. This system makes plants resistant to
CC the herbicide phenmedipham. Carbamate hydrolase has pH optimum 6.8, mol.
CC wt. 50-60kD, isoelectric point 6.2, and can cleave phenmedipham to methyl
CC 3-hydroxyphenylcarbamate, m-toluidine and CO2, so inactivating
CC it. It is produced by A. oxidans DSM 4044 which contains the 41 kb plasmid
CC pHp52.
XX
XX Sequence 493 AA;
```

Query Match 75.0%; Score 36; DB 10; Length 493;
Best Local Similarity 71.4%; Pred. No. 3.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 GGCVCVQD 8
IIIIII
DB 15 ggcvcvq 21

```
RESULT 11
G31664
ID G31664 standard; protein: 541 AA.
XX
XX
AC G31664;
XX
XX 17-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 38064.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP103405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
```

PR	14-MAY-1999;	9905-0134370.
PR	18-MAY-1999;	9905-0134768.
PR	18-MAY-1999;	9905-0134941.
PR	20-MAY-1999;	9905-0135124.
PR	21-MAY-1999;	9905-0135353.
PR	24-MAY-1999;	9905-0135622.
PR	25-MAY-1999;	9905-0136022.
PR	27-MAY-1999;	9905-0136392.
PR	28-MAY-1999;	9905-0136782.
PR	01-JUN-1999;	9905-0137222.
PR	03-JUN-1999;	9905-0137528.
PR	04-JUN-1999;	9905-0137502.
PR	07-JUN-1999;	9905-0137724.
PR	08-JUN-1999;	9905-0138094.
PR	10-JUN-1999;	9905-0138540.
PR	10-JUN-1999;	9905-0138847.
PR	14-JUN-1999;	9905-0139119.
PR	16-JUN-1999;	9905-0139452.
PR	17-JUN-1999;	9905-0139453.
PR	17-JUN-1999;	9905-0139492.
PR	18-JUN-1999;	9905-0139454.
PR	18-JUN-1999;	9905-0139455.
PR	18-JUN-1999;	9905-0139456.
PR	18-JUN-1999;	9905-0139457.
PR	18-JUN-1999;	9905-0139458.
PR	18-JUN-1999;	9905-0139459.
PR	18-JUN-1999;	9905-0139460.
PR	18-JUN-1999;	9905-0139461.
PR	18-JUN-1999;	9905-0139462.
PR	18-JUN-1999;	9905-0139463.
PR	18-JUN-1999;	9905-0139750.
PR	18-JUN-1999;	9905-0139751.
PR	21-JUN-1999;	9905-0139817.
PR	22-JUN-1999;	9905-0139859.
PR	23-JUN-1999;	9905-0140353.
PR	24-JUN-1999;	9905-0140354.
PR	28-JUN-1999;	9905-0140695.
PR	29-JUN-1999;	9905-0140823.
PR	30-JUN-1999;	9905-0140991.
PR	01-JUL-1999;	9905-0141287.
PR	01-JUL-1999;	9905-0141842.
PR	02-JUL-1999;	9905-0142154.
PR	06-JUL-1999;	9905-0142055.
PR	08-JUL-1999;	9905-0142390.
PR	09-JUL-1999;	9905-0142803.
PR	12-JUL-1999;	9905-0142920.
PR	13-JUL-1999;	9905-0142977.
PR	14-JUL-1999;	9905-0143542.
PR	15-JUL-1999;	9905-0143624.
PR	16-JUL-1999;	9905-0144005.
PR	16-JUL-1999;	9905-0144085.
PR	19-JUL-1999;	9905-0144086.
PR	19-JUL-1999;	9905-0144325.
PR	19-JUL-1999;	9905-0144331.
PR	19-JUL-1999;	9905-0144332.
PR	19-JUL-1999;	9905-0144333.
PR	19-JUL-1999;	9905-0144334.
PR	20-JUL-1999;	9905-0144335.
PR	20-JUL-1999;	9905-0144352.
PR	20-JUL-1999;	9905-0144632.
PR	20-JUL-1999;	9905-0144884.
PR	21-JUL-1999;	9905-0144884.
PR	21-JUL-1999;	9905-0145086.
PR	21-JUL-1999;	9905-0145088.
PR	22-JUL-1999;	9905-0145087.
PR	22-JUL-1999;	9905-0145087.
PR	22-JUL-1999;	9905-0145087.
PR	22-JUL-1999;	9905-0145088.
PR	22-JUL-1999;	9905-0145089.
PR	23-JUL-1999;	9905-0145145.
PR	23-JUL-1999;	9905-0145218.
PR	23-JUL-1999;	9905-0145224.
PR	26-JUL-1999;	9905-0145276.
PR	27-JUL-1999;	9905-0145913.

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 75.0%; Score 36; DB 21; Length 541;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCVCQ 6
Db 523 99gcvcq 528

RESULT 12
G31663
ID G31663 standard; Protein; 583 AA.
XX
AC G31663;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38063.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145210.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.

Query Match	75.0%	Score 36;	DB 19;	Length 1487;
Best Local Similarity	85.7%;	Pred. No. 1e+03;		
Matches	6;	Conservative	0;	Mismatches 1;
				Indels 0;
				Gaps 0;

OY 2 GGCVDG 8
 1 11111
 DB 32 gscvqpg 38

RESULT 14
 W88384 standard; Peptide: 31 AA.

AC W88384;

DT 26-APR-1999 (first entry)

DE Human Zneul EGF-like domain 1.

KW Zneul-1; neuro-growth factor-like protein; human; breast cancer;
 KW glioblastoma; pituitary adenoma; Alzheimer's disease; therapy;
 KW nerve regeneration; haematopoiesis; fertility; contraception;
 KW antibody; epidermal growth factor; EGF.

OS Homo sapiens.

PN W09857983-A2.

PD 23-DEC-1998.

PF 18-JUN-1998; 98MO-US12763.

PR 18-JUN-1997; 97US-0878322.

PR 18-JUN-1997; 97US-0050143.

PA (ZYMO) ZYMOGENETICS INC.

PI Blumberg H, Jelinek LJ, Lehner JM, Sheppard PO;

PI Whitmore TE;

DR WPI; 1999-095324/08.

PT New mammalian Zneul polypeptides - used to, e.g. treat Alzheimer's

PS disease, cancer and to repopulate blood cells

PS Claim 6; Page 53; 70pp; English.

CC This peptide comprises the first epidermal growth factor-like
 CC domain (EGF-1), i.e. amino acid residues 105-135, of human Zneul
 CC (see W88384), a new neuro-growth factor-like protein. Zneul can be
 CC used as a growth, maintenance, or differentiation factor in the
 CC spinal cord, heart, spleen, testis, thyroid and lymph nodes. It
 CC may also play a role in breast cancer, glioblastomas, and pituitary
 CC adenomas. Zneul may also be used to treat Alzheimer's disease,
 CC cancer, to repopulate blood cells after chemotherapy, to stimulate
 CC myofibroblast proliferation, stimulate or inhibit growth factors
 CC made in the placenta, in fertility and contraception, or to
 CC regenerate nerves. Claimed Zneul polypeptides (see also W88382-97),
 CC including specific domains of Zneul and epitope-bearing portions of
 CC Zneul, can be used to raise specific antibodies for use e.g. in
 CC diagnostic assays.

XX Sequence 31 AA;

Query Match 72.9%; Score 35; DB 20; Length 31;

Best Local Similarity 75.0%; Pred. No. 40;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGCVDG 8
 1 11111
 DB 10 gscvqpg 17

RESULT 15
 W88389 standard; Protein: 73 AA.

XX W88389;

DT 26-APR-1999 (first entry)

DE Human Zneul partial polypeptide.

KW Zneul-1; neuro-growth factor-like protein; human; breast cancer;
 KW glioblastoma; pituitary adenoma; Alzheimer's disease; therapy;
 KW nerve regeneration; haematopoiesis; fertility; contraception;
 KW antibody.

OS Homo sapiens.

PN W09857983-A2.

PD 23-DEC-1998.

PF 18-JUN-1998; 98MO-US12763.

PR 18-JUN-1997; 97US-0878322.

PR 18-JUN-1997; 97US-0050143.

PA (ZYMO) ZYMOGENETICS INC.

PI Blumberg H, Jelinek LJ, Lehner JM, Sheppard PO;

PI Whitmore TE;

DR WPI; 1999-095324/08.

PT New mammalian Zneul polypeptides - used to, e.g. treat Alzheimer's

PS disease, cancer and to repopulate blood cells

PS Claim 6; Page 57; 70pp; English.

CC This claimed polypeptide comprises amino acid residues 105-177 of
 CC the human Zneul mature protein (see W88382), i.e. the 2 epidermal
 CC growth factor-like domains (see also W88384-85) of Zneul. Zneul is
 CC a new neuro-growth factor-like protein that can be used as a
 CC growth, maintenance, or differentiation factor in the spinal cord,
 CC heart, spleen, testis, thyroid and lymph nodes. Zneul may also
 CC play a role in breast cancer, glioblastomas, and pituitary
 CC adenomas. Zneul can be used to treat Alzheimer's disease, cancer,
 CC to repopulate blood cells after chemotherapy, to stimulate
 CC myofibroblast proliferation, stimulate or inhibit growth factors
 CC made in the placenta, in fertility and contraception, or to
 CC regenerate nerves. Claimed Zneul polypeptides (see also W88382-97),
 CC including specific domains of Zneul and epitope-bearing portions of
 CC Zneul, can be used to raise specific antibodies for use e.g. in
 CC diagnostic assays.

XX Sequence 73 AA;

Query Match 72.9%; Score 35; DB 20; Length 73;

Best Local Similarity 75.0%; Pred. No. 89;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGCVDG 8
 1 11111
 DB 10 gscvqpg 17

Search completed: June 13, 2001, 14:25:44
 Job time: 663 sec

Wed Jun 13 15:00:17 2001

pct-us01-05825a-22.rag

Page 12

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:23:10 ; Search time 87.97 Seconds
(without alignments)
6.250 Million cell updates/sec

Title: PCT-US01-05825A-22

Perfect score: 48

Sequence: 1 GGCVCVQDG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR67:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	87.5	206	2	S18250 collagen alpha 1(I)
2	41	85.4	491	2	S56753 interferon regulat
3	39	81.2	196	2	F71525 hypothetical prote
4	39	81.2	196	2	H81681 maf protein TC0628
5	39	81.2	556	2	S22634 sphingomyelin phos
6	38	79.2	302	2	S59853 DNA-binding protel
7	38	79.2	2395	1	S50820 surface protein ty
8	37	77.1	279	2	G69001 endonuclease IV -
9	37	77.1	457	2	S39713 probable aldehyde
10	37	77.1	2704	2	S09118 G surface protein
11	37	77.1	2718	2	A23475 G surface protein
12	36	75.0	145	2	H69051 heterodisulfide re
13	36	75.0	153	2	A27179 collagen alpha 1(I)
14	36	75.0	362	2	A70547 probable mene - My
15	36	75.0	434	2	T20400 hypothetical prote
16	36	75.0	452	2	A83758 aldehyde dehydroge
17	36	75.0	493	2	A45737 phenylcarbamate hy
18	36	75.0	680	2	T30620 hypothetical prote
19	36	75.0	1193	2	T21133 hypothetical prote
20	36	75.0	1487	1	CGH06C collagen alpha 1(I)
21	36	75.0	1492	2	A40333 collagen alpha 1(I)
22	36	75.0	1548	2	S34583 serine proteinase
23	36	75.0	2233	2	T28669 surface protein 51
24	35	72.9	231	2	E69158 hypothetical prote
25	35	72.9	252	2	S08054 hypothetical prote
26	35	72.9	258	1	K1EC4E acetylglutamate Ki
27	34	70.8	296	2	A70973 hypothetical prote
28	34	70.8	447	2	T46835 precorrin-6y methy
29	34	70.8	618	2	S06446 citrolysin-related

30	34	70.8	633	2	T47524 hypothetical prote
31	34	70.8	680	1	XJBXTK transketolase (EC
32	34	70.8	681	2	S37809 transketolase (EC
33	34	70.8	685	2	T40162 transketolase - fl
34	34	70.8	695	2	T36007 probable transke
35	33.5	69.8	561	2	T27318 hypothetical prote
36	33.5	69.8	626	2	T27319 hypothetical prote
37	33	68.8	130	1	KRSH3A keratin high-sulfu
38	33	68.8	131	1	KRSH3 keratin high-sulfu
39	33	68.8	131	1	KRGT3M keratin high-sulfu
40	33	68.8	131	2	T05837 keratin high-sulfu
41	33	68.8	132	1	KRG73J keratin high-sulfu
42	33	68.8	141	2	I40577 hypothetical prote
43	33	68.8	161	2	A46189 vasotocin - Pacifi
44	33	68.8	188	2	H38192 chitin synthase (E
45	33	68.8	323	2	G83067 hypothetical prote

ALIGNMENTS

RESULT 1
S18250
collagen alpha 1(II) chain precursor - chicken (fragment)
C:Species: Gallus gallus (chicken)
C>Date: 22-Nov-1993 #sequence_revision 26-Jul-1996 #text_change 13-Aug-1999
C:Accession: S18250
R:NaH, H.D.; Upholt, W.B.
J. Biol. Chem. 266, 23446-23452, 1991
A:Title: Type II collagen mRNA containing an alternatively spliced exon predominates
A:Reference number: S18250; M0ID:92078225
A:Accession: S18250
A:Molecule type: mRNA
A:Residues: 1-206 <NAH>
A:Cross-references: EMBL:M74435; NID:g211635; PIDN:AAA48714.1; PID:g211636
A:Genetics:
C:Gene: COL2A1
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo
C:Keywords: alternative splicing; collid coll; extracellular matrix; glycoprotein
F:1-25/Domain: signal sequence #status predicted <Sto>
F:26-184/Domain: amino-terminal propeptide #status predicted <PRO>
F:36-95/Domain: von Willebrand factor type C repeat homology <WMC>
F:185-206/Product: collagen alpha 1(II) chain (fragment) #status predicted <MAT>

Query Match 87.5%; Score 42; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGCVCVQDG 8
DB 35 GGCVCVQDG 41

RESULT 2
S56753
Interferon regulatory factor 3 - chicken
C:Species: Gallus gallus (chicken)
C>Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jul-2000
A:Accession: S56753
R:Grant, C.E.; Vasa, M.Z.; Deeley, R.G.
Nucleic Acids Res. 23, 2137-2146, 1995.
A:Title: CIRF-3, a new member of the interferon regulatory factor (IRF) family that i
A:Reference number: S56753; M0ID:95334365
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-491 <CRA>
A:Cross-references: EMBL:U20338; NID:g790580; PIDN:AAA85995.1; PID:g790581
C:Superfamily: lymphoid-specific interferon regulatory factor

Query Match 85.4%; Score 41; DB 2; Length 491;

Best Local Similarity 87.5%; Pred. No. 9.8;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCVODG 8
Db 235 GGGCGQDG 242

RESULT 3

F71525

hypothetical protein CT349 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C:Accession: F71525

R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998

A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
A:Reference number: A71570; MUID:99000809

A:Accession: F71525
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-196 <ARN>
A:Cross-references: GB:AE001308; GB:AE001273; NID:93328766; PIDN:AA67944.1; PID:9332877

A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: CT349

C:Superfamily: septum formation protein maf

Query Match
Best Local Similarity 81.2%; Score 39; DB 2; Length 196;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCVODG 8
Db 150 GGGCVODG 157

RESULT 4

H81681

maf protein TC0628 [Imported] - Chlamydia muridarum (strain N199)
C:Species: Chlamydia muridarum, Chlamydia trachomatis Mopn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255

A:Accession: H81681
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-196 <GEN>
A:Cross-references: GB:AE002331; GB:AE002160; NID:97190662; PIDN:AAF39457.1; PID:9719066

A:Experimental source: strain N199 (Mopn)
C:Genetics:
A:Gene: TC0628

C:Superfamily: septum formation protein maf

Query Match
Best Local Similarity 81.2%; Score 39; DB 2; Length 196;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCVODG 8
Db 150 GGGCVODG 157

RESULT 5

S22634

sphingomyelin phosphodiesterase (EC 3.1.4.12) precursor - Leptospira interrogans
N:Alternate names: sphingomyelinase C

C:Species: Leptospira interrogans
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 08-Oct-1999
C:Accession: S22634; S10466
R:Segers, R.P.A.M.; van der Drift, A.; de Nijls, A.; Corcione, P.; van der Zeijst, B.A.
Infect. Immun. 58, 2177-2185, 1990
A:Title: Molecular analysis of a sphingomyelinase C gene from Leptospira interrogans
A:Reference number: S22634; MUID:90507220
A:Accession: S22634
A:Molecule type: DNA
A:Residues: 1-556 <SEG>
A:Cross-references: EMBL:X52176; NID:944012; PIDN:CAA36424.1; PID:944013
C:Keywords: phosphoric diester hydrolase
F:1-27/Domain: signal sequence #status predicted <Sig>
F:28-556/Product: sphingomyelin phosphodiesterase #status predicted <Mat>

Query Match
Best Local Similarity 81.2%; Score 39; DB 2; Length 556;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 GGGCVODG 8
Db 488 GGGCVODG 494

RESULT 6

S59853

DNA-binding protein Dof3 - maize (fragment)
C:Species: Zea mays (maize)
C:Date: 15-Feb-1996 #sequence_revision 19-Apr-1996 #text_change 29-Oct-1999

R:Spunich, Y.
submitted to the EMBL Data Library, June 1994

A:Reference number: S59852
A:Accession: S59853
A:Molecule type: mRNA

A:Residues: 1-302 <SHU>
A:Cross-references: EMBL:X79935; NID:91061307; PIDN:CAA56288.1; PID:91061308

A:Title: A novel DNA-binding domain that may form a single zinc finger motif.
A:Reference number: S59850; MUID:96032831

A:Accession: S59851
A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA
A:Residues: 27-99 <YAN>
A:Cross-references: EMBL:X79935

C:Keywords: zinc finger

Query Match
Best Local Similarity 79.2%; Score 38; DB 2; Length 302;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCVODG 8
Db 240 GGGCGEDG 247

RESULT 7

S50820

surface protein type 51B - Paramoecium tetraurelia
C:Species: Paramoecium tetraurelia
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 07-Dec-1999

R:Scott, J.; Iseck, C.; Forney, J.
Nucleic Acids Res. 22, 5079-5084, 1994

A:Title: Analysis of the micronuclear B type surface protein gene in Paramoecium tetra
A:Reference number: S50820; MUID:95098630
A:Accession: S50820

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2395 <SCO>

A:Cross-references: EMBL:U07603; NID:g467226; PIDN:AA81947.1; PID:g467227
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1994
 C:Genetics:
 A:Genetic code: SGC5
 A:introns: 472/3; 1310/3; 1821/3
 C:Superfamily: G surface protein

Query Match 79.2% Score 38; DB 1; Length 2395;
 Best Local Similarity 75.0% Pred. No. 1.4e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCVDG 8
 |||||:1
 Db 327 GGGCVSNG 334

RESULT 8
 G69001
 endonuclease IV - Methanobacterium thermoautotrophicum (strain Delta H)
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 08-Oct-1999
 C:Accession: G69001
 R:Smith, D.R.; Doucette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 Olm, D.; Spadofora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.;
 K. S.; Church, G.M.; Dajile, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
 A:Reference number: A69000; MUID:98037514
 A:Accession: G69001
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-279 <MTH>
 A:Cross-references: GB:AEO00874; GB:AEO00666; NID:g2622110; PIDN:AA85506.1; PID:g262211
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH1010
 A:Start codon: GTG
 C:Superfamily: Methanobacterium thermoautotrophicum endonuclease IV

Query Match 77.1% Score 37; DB 2; Length 279;
 Best Local Similarity 71.4% Pred. No. 29;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCVD 7
 |||||:1
 Db 185 GGGCIRD 191

RESULT 9
 S39713
 Probable aldehyde dehydrogenase (NAD+) (EC 1.2.1.3) ywdH - Bacillus subtilis
 N:Alternate names: protein lpa-58r
 C:Species: Bacillus subtilis
 C:Date: 07-Oct-1994 #sequence_revision 26-May-1995 #text_change 20-Jun-2000
 C:Accession: S39713; C70054
 R:Glaser, P.; Kunst, F.; Arnaut, M.; Coudart, M.P.; Gonzales, W.; Hullo, M.F.; Ionescu,
 A.; Rapoport, G.; Danchin, A.
 Mol. Microbiol. 10, 371-384, 1993
 A:Title: Bacillus subtilis genome project: cloning and sequencing of the 97 kb region fr
 A:Reference number: S39655; MUID:95020537
 A:Accession: S39713
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-457 <GLA>
 A:Cross-references: EMBL:X73124; NID:g413923; PIDN:CAA51614.1; PID:g413982
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1993
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beret
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
 A.; Ehrlich, S.D.; Emmerich, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzl, A.; Galle

lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.
 koetter, P.; Koningstein, G.; Krogan, S.; Kumano, M.; Kurita, K.; Lapius, A.; Lardino
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Ma
 Y., M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanl
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchida
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili
 A:Reference number: A69580; MUID:98044033
 A:Accession: C70054
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-457 <KUN>
 A:Cross-references: GB:Z99123; GB:AL009126; NID:g2636240; PIDN:CAB15822.1; PID:g26363
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: ywdH
 C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology
 C:Keywords: oxidoreductase
 F:4-265/Domain: aldehyde dehydrogenase homology <ALDD>

Query Match 77.1% Score 37; DB 2; Length 457;
 Best Local Similarity 85.7% Pred. No. 45;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCVD 7
 |||||:1
 Db 385 GGGCVD 391

RESULT 10
 S09118
 G surface protein 168 - Paramecium primaurelia
 C:Species: Paramecium primaurelia
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 07-Dec-1999
 C:Accession: S09118
 R:Prat, A.
 J. Mol. Biol. 211, 521-535, 1990
 A:Title: Conserved sequences flank variable tandem repeats in two alleles of the G su
 A:Reference number: S09118; MUID:90172419
 A:Accession: S09118
 A:Molecule type: DNA
 A:Residues: 1-2704 <PRA>
 A:Cross-references: EMBL:X52133; NID:g10049; PIDN:CAA36378.1; PID:g578473
 C:Genetics:
 A:Genetic code: SGC5
 C:Superfamily: G surface protein

Query Match 77.1% Score 37; DB 2; Length 2704;
 Best Local Similarity 75.0% Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGCVDG 8
 |||||:1
 Db 189 GGGCVDG 196

RESULT 11
 A23475
 G surface protein - Paramecium primaurelia
 C:Species: Paramecium primaurelia
 C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 07-Dec-1999
 C:Accession: A23475
 R:Prat, A.; Katinka, M.; Caron, F.; Meyer, E.
 J. Mol. Biol. 189, 47-60, 1986
 A:Title: Nucleotide sequence of the Paramecium primaurelia G surface protein. A huge
 A:Reference number: A23475; MUID:87060934
 A:Accession: A23475
 A:Molecule type: DNA

A:Residues: 1-2718 <PRA>
 A>Note: The authors translated the codon TGC for residue 2665 as Trp
 C:Genetics:
 C:Genetic code: SCCS
 C:Superfamily: G surface protein

Query Match 77.1%; Score 37; DB 2; Length 2718;
 Best Local Similarity 75.0%; Pred. No. 2.3e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GGCVCVQDG 8
 |||||
 Db 195 GGCVCVQDG 202

RESULT 12
 H69051
 heterodisulfide reductase HdrD related protein - Methanobacterium thermoautotrophicum (S
 C:Species: Methanobacterium thermoautotrophicum
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: H69051
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 ; Qiu, D.; Spadotora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
 J. S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A>Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
 A:Reference number: A69000; MUID:98037514
 A:Accession: H69051
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-145 <MTH>
 A:Cross-references: GB:AE000803; GB:AE000666; NID:g2621179; PIDN:AA84645.1; PID:g262118
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH139

Query Match 75.0%; Score 36; DB 2; Length 145;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGCVCVQ 6
 |||||
 Db 100 GGCVCVQ 105

RESULT 13
 A27179
 collagen alpha 1(I) chain precursor - chicken (fragment)
 C:Species: Gallus gallus (chicken)
 C>Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 13-Aug-1999
 C:Accession: A27179; A29367
 R:Flner, M.H.; Boedtker, H.; Doty, P.
 Gene 56, 71-78, 1987
 A>Title: Construction and characterization of cDNA clones encoding the 5' end of the ch
 A:Reference number: A27179; MUID:88056316
 A:Accession: A27179
 A:Molecule type: mRNA
 A:Residues: 1-153 <FTN>
 A:Cross-references: GB:M17607; NID:g211473; PIDN:AAA8672.1; PID:g211474
 R:Flner, M.H.; Aho, S.; Gerstenfeld, L.C.; Boedtker, H.; Doty, P.
 J. Biol. Chem. 263, 13323-13332, 1987
 A>Title: Unusual DNA sequences located within the promoter region and the first intron c
 A:Reference number: A29367; MUID:88007542
 A:Accession: A29367
 A:Molecule type: DNA
 A:Residues: 1-144 <FT2>
 A:Cross-references: GB:M17607
 C:Genetics:
 A:Introns: 28/1; 93/1; 104/3; 116/3
 C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxyllysine; hydroxyprol

F:1-22/Domain: signal sequence #status predicted <SIG>
 F:23-153/Domain: amino-terminal propeptide #status predicted <PRO>
 F:32-91/Domain: von Willebrand factor type C repeat homology <VWC>

Query Match 75.0%; Score 36; DB 2; Length 153;
 Best Local Similarity 85.7%; Pred. No. 25;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 GGCVCVQDG 8
 |||||
 Db 31 GGCVCVQDG 37

RESULT 14
 A70547
 probable mene - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C:Accession: A70547
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
 ; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holtroyd,
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
 A:Reference number: A70500; MUID:98295987
 A:Accession: A70547
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-362 <COL>
 A:Cross-references: GB:Z95558; GB:AL123456; NID:g3261781; PIDN:CAB08976.1; PID:g31686
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: mene

Query Match 75.0%; Score 36; DB 2; Length 362;
 Best Local Similarity 85.7%; Pred. No. 54;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 GGCVCVQDG 8
 |||||
 Db 185 GGCVCVQDG 191

RESULT 15
 T20400
 hypothetical protein E01G4.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
 C:Accession: T20400
 R:Sims, M.
 submitted to the EMBL Data Library, December 1996
 A:Reference number: Z19269
 A:Accession: T20400
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-434 <WIL>
 A:Cross-references: EMBL:Z83223; PIDN:CAB05714.1; GSPDB:GN00020; CESP:E01G4.2
 A:Experimental source: clone E01G4
 C:Genetics:
 A:Gene: CESP:E01G4.2
 A:Map position: 2
 A:Introns: 47/3; 69/3; 105/3; 224/3; 400/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein E01G4.2

Query Match 75.0%; Score 36; DB 2; Length 434;
 Best Local Similarity 85.7%; Pred. No. 63;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGCVCVQ 7

|||||
Db 332 GGCVAD 338

Search completed: June 13, 2001, 14:23:10
Job time: 741 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:30:35 ; Search time 51.57 Seconds
(Without alignments)
5.314 Million cell updates/sec

Title: PCT-US01-05825A-22
Perfect score: 48
Sequence: 1 GGGCVDG 8

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	85.4	491	1	IRF3_CHICK
2	39	81.2	556	1	PHL_LEPIN
3	37	77.1	457	1	DHA2_BACSU
4	37	77.1	2704	1	G168_PARR
5	37	77.1	2715	1	G156_PARR
6	36	75.0	493	1	PCD_ARTOX
7	36	75.0	1453	1	CA11_CHICK
8	36	75.0	1877	1	PCF5_MOUSE
9	35	72.9	256	1	YMS4_MAIZE
10	35	72.9	258	1	ARGB_ECOLI
11	34	70.8	343	1	GLGD_BACST
12	34	70.8	618	1	CIR1_CITFR
13	34	70.8	621	1	YC92_CAEBL
14	34	70.8	679	1	TKT1_KLULA
15	34	70.8	679	1	TKT1_YEAST
16	34	70.8	681	1	TKT2_YEAST
17	34	70.8	695	1	TKT_YEAST
18	33	68.8	130	1	KRA3_SHEEP
19	33	68.8	131	1	KRA3_SHEEP
20	33	68.8	132	1	KRA3_CAPHI
21	33	68.8	189	1	CHS1_ASPNG
22	33	68.8	189	1	CHS1_EMENT
23	33	68.8	289	1	SCRK_LACLC
24	33	68.8	394	1	HYFL_ALCEU
25	33	68.8	551	1	ERR_MOUSE
26	33	68.8	557	1	TCPA_DROME
27	33	68.8	559	1	TCPA_DROME
28	33	68.8	561	1	EAT4_MOUSE
29	33	68.8	561	1	EAT4_RAT
30	33	68.8	587	1	URE1_CLOPE
31	33	68.8	625	1	DGD_CAEBL
32	33	68.8	640	1	CLAT_PIG
33	33	68.8	2318	1	NTC3_MOUSE

34	32	66.7	150	1	THH_MOUSE	062264 mus musculus
35	32	66.7	150	1	THH_RAT	P04143 rattus norv
36	32	66.7	223	1	RB21_CANFA	P55745 canis famill
37	32	66.7	247	1	CAH_METTE	P40881 methanosarc
38	32	66.7	278	1	EUTJ_ECOLI	P77277 escherichia
39	32	66.7	279	1	EUTJ_SALTY	P41794 salmoeella
40	32	66.7	316	1	CH18_POPTR	P16061 populus trl
41	32	66.7	340	1	ASPG_FLAME	047898 flavobacter
42	32	66.7	451	1	Y996_METUA	058408 methanococc
43	32	66.7	473	1	LE02_MYCTU	053237 mycobacteri
44	32	66.7	481	1	LORI_MOUSE	P18165 mus musculu
45	32	66.7	485	1	SSGP_VOLCA	P21997 volvox cart

ALIGNMENTS

```

RESULT 1
ID IRF3_CHICK STANDARD; PRT: 491 AA.
AC 090643;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE INTERFERON REGULATORY FACTOR 3 (IRF-3).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=95334365; PubMed=7541908;
RX Grant C.E., Vasa M.Z., Deeley R.G.;
RT "IRF-3, a new member of the interferon regulatory factor (IRF)
RT family that is rapidly and transiently induced by dsRNA.";
RL Nucleic Acids Res. 23:2137-2146(1995).
CC -!- FUNCTION: MAY ACTIVATE TRANSCRIPTION BY COMPLEX FORMATION WITH
CC OTHER TRANSCRIPTIONAL FACTORS, POSSIBLY MEMBERS OF THE STAT
CC FAMILY. BINDS SPECIFICALLY TO THE IFN-STIMULATED RESPONSE ELEMENT
CC (ISRE) BUT NOT TO THE IRF-1 BINDING SITE PRD-1.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U20338; AACB6995.1; -
CC HSSP: P15314; IRF1
CC InterPro: IPR001346; -
CC Pfam: PF00605; IRF_1
CC PRINTS: PR00267; INTERFERGCT.
CC PROSITE: PS00601; IRF_1
CC Transcription regulation; DNA-binding; Nuclear protein; Activator.
FT DNA_BIND 14 112 TRYPTOPHAN PENNAD REPEAT.
SQ SEQUENCE 491 AA: 54441 MW: CAE0C2AABE976D9 CRC64;

```

Query Match 85.4%; Score 41; DB 1; Length 491;

Best Local Similarity 87.5%; Pred. No. 4.2; Mismatches 1; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 235 GGGCVDG 242

RESULT 2

PHL_LEPIN ID STANDARD: PRT: 556 AA.
 AC P17627;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE SPHINGOMYELINASE C PRECURSOR (EC 3.1.4.12) (SPHINGOMYELIN
 DE PHOSPHODIESTERASE) (SMASE).
 OS Leptospira interrogans.
 OC Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
 NCBI_TaxId=173;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SEVOVAR HARDJO / ISOLATE SPONSELEE;
 RA MEDLINE=90307220; PubMed=2163985;
 RA Segers R.P.A.M., van der Drift A., de Nijls A., Corcione P.,
 RA van der Zeijst B.A.M., Gaastra W.;
 RT "Molecular analysis of a sphingomyelinase C gene from Leptospira
 RT interrogans serovar hardjo".
 RL Infect. Immun. 58:2177-2185(1990).
 CC -1- CATALYTIC ACTIVITY: SPHINGOMYELIN + H(2)O = N-ACYLSPHINGOSINE +
 CC CHOLINE PHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: TO BACILLUS SPHINGOMYELINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X52176; CAA36424.1; -
 DR PIR: S22634; S22634.
 KW Hydrolyase; Signal.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 556 SPHINGOMYELINASE C.
 SO SEQUENCE 556 AA; 63268 MW; 6D0EACDB9A5CD6FD CRC64;

Query Match 81.2%; Score 39; DB 1; Length 556;
 Best Local Similarity 85.7%; Pred. No. 11;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GGCYVDG 8
 DB 488 GGCYVDG 494

RESULT 3
 DHM2_BACSU STANDARD: PRT: 457 AA.
 ID AC P39616;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 13-JUL-1998 (Rel. 36, Last annotation update)
 DE PROBABLE ALDEHYDE DEHYDROGENASE YMDH (EC 1.2.1.3).
 GN YMDH OR IPA-58R.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Bacillus.
 NCBI_TaxId=1423;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA MEDLINE=95020537; PubMed=7934828;
 RA Glaeser P., Kunst F., Arnold M., Coudart M.P., Gonzales W.,
 RA Hullo M.F., Ionescu W., Lubochinsky B., Marcelino L., Moszer I.,
 RA Presecan E., Santana M., Schneider E., Schweizer J., Vertes A.,
 RA Rapoport G., Danchin A.;
 RT "Bacillus subtilis genome project: cloning and sequencing of the 97

RT kb region from 325 degrees to 333 degrees.*;
 RL Mol. Microbiol. 10:371-384(1993).
 CC -1- CATALYTIC ACTIVITY: ALDEHYDE + NAD(+) + H(2)O = ACID + NADH.
 CC -1- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: X73124; CAA51614.1; -
 DR EMBL: 299123; CAB15822.1; -
 DR PIR: S39713; S39713.
 DR HSSP: P11883; 1ND3.
 DR Subtilisin; B610604; ywdH.
 DR InterPro: IPR002086; -
 DR Pfam: PF00171; aldedh; 1.
 DR PROSITE: PS00070; ALDEHYDE DEHYDR. CYS. 1.
 DR PROSITE: PS00687; ALDEHYDE DEHYDR. GLU. 1.
 KW Hypothetical protein; Oxidoreductase; NAD.
 FT ACT_SITE 211 211 BY SIMILARITY.
 FT ACT_SITE 245 245 BY SIMILARITY.
 SO SEQUENCE 457 AA; 50765 MW; C238D9FD07DFB92A CRC64;

Query Match 77.1%; Score 37; DB 1; Length 457;
 Best Local Similarity 85.7%; Pred. No. 19;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGCYVD 7
 DB 385 GGCYVD 391

RESULT 4
 G168_PAPR STANDARD: PRT: 2704 AA.
 ID AC P17053;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-AUG-1990 (Rel. 15, Last annotation update)
 DE 1686 SURFACE PROTEIN PRECURSOR.
 GN 1686.
 OS Paramacium primaurelia.
 OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
 CC Paramacium.
 NCBI_TaxId=5886;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA MEDLINE=90172419; PubMed=2308165;
 RA Prat A.;
 RT "Conserved sequences flank variable tandem repeats in two alleles of
 RT the G surface protein of Paramacium primaurelia.";
 RL J. Mol. Biol. 211:521-535(1990).
 CC -1- FUNCTION: THIS PROTEIN IS THE SURFACE ANTIGEN OR IMMOBILIZATION
 CC ANTIGEN OF PARAMACIUM PRIMAURELIA.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
 CC -1- MISCELLANEOUS: IT HAS INTERNAL HOMOLOGIES AND A HIGHLY PERIODIC
 CC STRUCTURE WITH 37 PERIODS OF ABOUT 75 RESIDUES, EACH PERIOD
 CC CONTAINING 8 CYSTEINES, EXCEPT FOR FOUR HALF PERIODS. A VARIABLE
 CC PART OF 475 RESIDUES COMPRISES 4 ALMOST IDENTICAL PERIODS IN THE
 CC MIDDLE OF THE PROTEIN.
 CC -1- MISCELLANEOUS: EXPRESSION OF G PROTEIN OCCURS AT LOW TEMPERATURES
 CC (14-32 DEGREES CELSIUS).
 CC -1- SIMILARITY: 98% TO THE ALLELIC FORM 156G PROTEIN (P13837) IN
 CC PERIODIC STRUCTURE AND 80% IN THE VARIABLE DOMAIN IN THE MIDDLE
 CC OF THE PROTEIN.
 CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X52133; CAA36378.1; -.
DR PIR: S09118; S09118.
DR HSSP: P06620; IINA.
DR InterPro: IPR002895; -.
DR Pfam: PF01508; Paramedium_SA: 33.
KW Signal: Repeat; Antigen; Membrane; GPI-anchor.
FT SIGNAL 1 20
FT CHAIN 21 2704 1686 SURFACE PROTEIN.
FT DOMAIN 106 2560 37 X 75 AA APPROXIMATE REPEATS.
FT DOMAIN 1060 1424 VARIABLE DOMAIN, COMPRISES 4 ALMOST
FT IDENTICAL REPEATS.
SQ SEQUENCE 2704 AA; 278775 MW; 40E9A0B18EE2119 CRC64;

Query Match 77.1%; Score 37; DB 1; Length 2704;
Best Local Similarity 75.0%; Pred. No. 99;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGCYQDG 8
Db 189 GGCYDSC 196

RESULT 5
ID G156_PARRP STANDARD: PRT; 2715 AA.
AC P13837;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-AUG-1990 (Rel. 15, Last annotation update)
DE 156G SURFACE PROTEIN PRECURSOR.
GN 156G.
OS Paramedium primaurella.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Penicillida;
OC Paramedium.
OX NCBI_Taxid=5886;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=156;
RX MEDLINE=87060934; Pubmed=3783679;
RA Pratt A., Katinka M., Caron F., Meyer E.;
RT "Nucleotide sequence of the Paramedium primaurella G surface protein.
RT A huge protein with a highly periodic structure.";
RL J. Mol. Biol. 189:47-60(1986).
CC -1- FUNCTION: THIS PROTEIN IS THE SURFACE ANTIGEN OR IMMOBILIZATION
CC ANTIGEN OF PARAMEDIUM PRIMAURELLA.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
CC -1- MISCELLANEOUS: IT HAS INTERNAL HOMOLOGIES AND A HIGHLY PERIODIC
CC STRUCTURE WITH 37 PERIODS OF ABOUT 75 RESIDUES, EACH PERIOD
CC CONTAINING 8 CYSTEINES, EXCEPT FOR FOUR HALF PERIODS. A VARIABLE
CC PART OF 475 RESIDUES COMPRISES 4 ALMOST IDENTICAL PERIODS IN THE
CC MIDDLE OF THE PROTEIN.
CC -1- MISCELLANEOUS: EXPRESSION OF G PROTEIN OCCURS AT LOW TEMPERATURES
CC (14-32 DEGREES CELSIUS).
CC -1- SIMILARITY: 98% TO THE ALLELIC FORM 168G PROTEIN (P17053) IN
CC PERIODIC STRUCTURE AND 80% IN VARIABLE DOMAIN IN THE MIDDLE OF
CC THE PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: X03882; CAA27514.1; -.
DR PIR: A23475; A23475.
DR HSSP: P06620; IINA.
DR InterPro: IPR002895; -.
DR Pfam: PF01508; Paramedium_SA: 33.
KW Signal: Repeat; Antigen; Membrane; GPI-anchor.
FT SIGNAL 1 20
FT CHAIN 21 2715 156G SURFACE PROTEIN.
FT DOMAIN 106 2560 37 X 75 AA APPROXIMATE REPEATS.
FT SIMILAR 1 222 88% TO PARAMEDIUM TETRAURELLA A
FT PROTEIN.
SQ SEQUENCE 2715 AA; 279551 MW; 97BE355AB9C7C298 CRC64;

Query Match 77.1%; Score 37; DB 1; Length 2715;
Best Local Similarity 75.0%; Pred. No. 99;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGCYQDG 8
Db 195 GGCYDSC 202

RESULT 6
ID PCD_ARTOX STANDARD: PRT; 493 AA.
AC Q01470;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE PHENMEDIPHAM HYDROLASE (EC 3.1.1.-) (PHENYLCARBAMATE HYDROLASE).
GN PCD.
OS Arthrobacter oxidans.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Bacilli; Bacillales; Micrococcales; Micrococcaceae; Arthrobacter.
OX NCBI_Taxid=1671;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 264-270; 272-281 AND 410-427.
RC STRAIN=P52;
RX MEDLINE=93015712; Pubmed=1400211;
RA Pohlenz H.D., Boidol W., Schuetzke I., Streber W.R.;
RT "Purification and properties of an Arthrobacter oxydans P52 carbamate
RT hydrolase specific for the herbicide phenmedipham and nucleotide
RT sequence of the corresponding gene.";
RL J. Bacteriol. 174:6600-6607(1992).
CC -1- FUNCTION: MAY DEGRADE THE PHENYLCARBAMATE HERBICIDES PHENMEDIPHAM
CC AND DESMEDIPHAM COMETABOLICALLY BY HYDROLYZING THEIR CENTRAL
CC CARBAMATE LINKAGES. CONVEYS RESISTANCE TO THE HERBICIDE
CC PHENMEDIPHAM.
CC -1- SUBUNIT: MONOMER.
CC -1- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYL ESTERASE/LIPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M94965; AAA22078.1; -.
DR EMBL: A12923; CAA01069.1; -.
DR PIR: S27481; S27481.
DR PIR: A45737; A45737.
DR HSSP: P21836; IMAH.
DR InterPro: IPR002018; -.
DR Pfam: PF00135; Coesterase_1.
DR PROSITE: PS00122; CARBOXYL ESTERASE_B_1; 1.
DR PROSITE: PS00941; CARBOXYL ESTERASE_B_2; 1.
KW Hydrolase; Serine esterase; Herbicide resistance; Plasmid.
FT ACT SITE 188 188 BY SIMILARITY.
SQ SEQUENCE 493 AA; 53728 MW; 83CECE7F3194BB66 CRC64;

Query Match 75.0% Score 36; DB 1; Length 493;
 Best Local Similarity 71.4% Pred. No. 31;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 2 GGCYVDG 8
 DB 15 GGCLEDG 21

RESULT 7
 CALL_CHICK STANDARD: PRT: 1453 AA.
 AC P02457;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE COLLAGEN ALPHA 1(I) CHAIN PRECURSOR.
 GN COL1A1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE OF 1-153 FROM N.A.
 RA MEDLINE=86056316; PubMed=3678834;
 RT Finer M.H., Boedtker H., Doty P.;
 RT "Construction and characterization of cDNA clones encoding the 5' end
 of the chicken pro alpha 1(I) collagen mRNA.";
 RL Gene 56:71-78(1987).
 RN [2]
 RP SEQUENCE OF 1-144 FROM N.A.
 RA MEDLINE=88007542; PubMed=2820966;
 RT Finer M.H., Aho S., Gerstenfeld L.C., Boedtker H., Doty P.;
 RT "Unusual DNA sequences located within the promoter region and the
 first intron of the chicken pro-alpha 1(I) collagen gene.";
 RL J. Biol. Chem. 262:13323-13332(1987).
 RN [3]
 RP SEQUENCE OF 152-1187.
 RA MEDLINE=82231995; PubMed=7093229;
 RT Hahnerberger J.H., Corbett C., Dixit S.N., Yu W., Seyer J.M.,
 RA Kang A.H., Gross J.;
 RT "Amino acid sequence of chick skin collagen alpha 1(I)-C88 and the
 complete primary structure of the helical portion of the chick skin
 collagen alpha 1(I) chain.";
 RL Biochemistry 21:2048-2055(1982).
 RN [4]
 RP SEQUENCE OF 1200-1205.
 RA MEDLINE=72243016; PubMed=5047697;
 RT Eyre D.R., Glimcher M.J.;
 RT "Evidence for a previously undetected sequence at the carboxyterminus
 of the alpha 1 chain of chicken bone collagen";
 RL Biochem. Biophys. Res. Commun. 48:720-726(1972).
 RN [5]
 RP SEQUENCE OF 981-1453 FROM N.A.
 RA MEDLINE=81160715; PubMed=6927845;
 RT Fuller F., Boedtker H.;
 RT "Sequence determination and analysis of the 3' region of chicken pro-
 alpha 1(I) and pro-alpha 2(I) collagen messenger ribonucleic acids
 including the carboxy-terminal propeptide sequences.";
 RL Biochemistry 20:996-1006(1981).
 RN [6]
 RP SEQUENCE OF 1311-1453 FROM N.A.
 RA MEDLINE=80134546; PubMed=6987088;
 RT Showalter A.M., Pesciotta D.M., Eikenberry E.F., Yamamoto T.,
 RA Pastan I., Decrombrughe B., Fietzek P.P., Olsen B.R.;
 RT "Nucleotide sequence of a collagen cDNA-fragment coding for the
 carboxyl end of pro alpha 1(I)-chains";
 RL FEBS Lett. 111:61-65(1980).
 CC -1- FUNCTION: TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
 (FIBRILLAR FORMING COLLAGEN).

CC -1- SUBUNIT: TRIMERS OF ONE ALPHA 2(I) AND TWO ALPHA 1(I) CHAINS.
 CC -1- TISSUE SPECIFICITY: FORMS THE FIBRILS OF TENDON LIGAMENTS AND
 CC BONES. IN BONES THE FIBRILS ARE MINERALIZED WITH CALCIUM
 CC HYDROXYAPATITE.
 CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -1- SIMILARITY: CONTAINS 1 VMFC DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M17839; AAA48704.1;
 DR EMBL: M17838; AAA48704.1; JOINED.
 DR EMBL: V00401; CAA23695.1;
 DR EMBL: M10571; AAA48671.1; ALT_SEQ.
 DR EMBL: M17607; AAA48672.1;
 DR PIR: A02857; CGCHLS.
 DR PIR: A27179; A27179.
 DR PIR: A29367; A29367.
 DR InterPro: IPR000087;
 DR InterPro: IPR000885;
 DR InterPro: IPR01007;
 DR Pfam: PF01410; COLFI.1.
 DR Pfam: PF01391; Collagen.18.
 DR Pfam: PF00093; vwc.1.
 DR PROSITE: PS01208; VMFC.1.
 DR Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Glycoprotein; Collagen; Signal.
 FT SIGNAL 1 22
 FT PROPEP 23 151
 FT CHAIN 152 1205
 FT PROPEP 1206 1453
 FT DOMAIN 31 89
 FT MOD_RES 152 152
 FT MOD_RES 254 254
 FT MOD_RES 851 851
 FT MOD_RES 1081 1081
 FT MOD_RES 1097 1097
 FT MOD_RES 1153 1153
 FT CONFLICT 1187 1187
 FT CONFLICT 1441 1441
 FT CONFLICT 1441 1441
 SQ SEQUENCE 1453 AA; 137789 MW; 3BC6152134271F4D CRC64;
 OY 2 GGCYVDG 8
 DB 31 GGCYVDG 37

Query Match 75.0% Score 36; DB 1; Length 1453;
 Best Local Similarity 85.7% Pred. No. 83;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 2 GGCYVDG 8
 DB 31 GGCYVDG 37

RESULT 8
 PCK5_MOUSE STANDARD: PRT: 1877 AA.
 AC Q04592; G62040;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PROPROTEIN CONVERTASE SUBTILISIN/KEXIN TYPE 5 PRECURSOR (EC 3.4.21.-)
 DE (PROPROTEIN CONVERTASE PCS) (SUBTILISIN/KEXIN-LIKE PROTEASE PCS)
 DE (CONVERTASE PCS) (PC6) (SUBTILISIN-LIKE PROPROTEIN CONVERTASE 6)
 DE (SPC6).
 GN PCK5.
 OS Mus musculus (Mouse).

CC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN (1)
RN SEQUENCE OF 330-1877 FROM N.A. (LONG ISOFORM).
RP STRAIN-ICR; TISSUE=Intestine.
RC MEDLINE=93327933; PubMed=8335106;
RX Nakagawa T., Murakami K., Nakayama K.;
RA Nakagawa T., Hosaka M., Torii S., Watanabe T., Murakami K.,
RT "Identification and functional expression of a new member of the
RT mammalian kex2-like processing endoprotease family: its striking
RT structural similarity to PACE4.";
RL J. Biochem. 113:132-135(1993).
RN (3)
RN SEQUENCE FROM N.A. (SHORT ISOFORM).
RP TISSUE=Brain, and Intestine.
RC MEDLINE=931224489; PubMed=8468318;
RX Nakagawa T., Hosaka M., Torii S., Watanabe T., Murakami K.,
RT "Identification and functional expression of a new member of the
RT mammalian kex2-like processing endoprotease family: its striking
RT structural similarity to PACE4.";
RL J. Biochem. 113:132-135(1993).
RN (3)
RN SEQUENCE FROM N.A. (SHORT ISOFORM).
RP TISSUE=Adrenal cortex.
RC MEDLINE=93342056; PubMed=8311687;
RX Lussan J., Vieu D., Hamelin J., Day R., Chretien M., Seidah N.G.;
RA "cDNA structure of the mouse and rat subtilisin/kexin-like PCS5: a
RT candidate proprotein convertase expressed in endocrine and
RT nonendocrine cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:6691-6695(1993).
RN (4)
RN PARTIAL SEQUENCE, AND SUBCELLULAR LOCATION.
RP MEDLINE=97103178; PubMed=8947550;
RX De Bie I., Marchinkiewicz M., Malile D., Lazure C., Nakayama K.,
RA Bendayan M., Seidah N.G.;
RT "The isoforms of proprotein convertase PCS5 are sorted to different
RT subcellular compartments.";
RL J. Cell Biol. 135:1261-1275(1996).
RN (5)
RN DEVELOPMENTAL EXPRESSION.
RP MEDLINE=96293359; PubMed=8698813;
RX Constant D.B., Calton M., Robertson E.J.;
RA "SPC4, SPC6, and the novel protease SPC7 are coexpressed with bone
RT morphogenetic proteins at distinct sites during embryogenesis.";
RL J. Cell Biol. 134:181-191(1996).
RN (6)
RN DEVELOPMENTAL EXPRESSION.
RP MEDLINE=97436919; PubMed=9291583;
RX Rancourt S.L., Rancourt D.E.;
RA "Murine subtilisin-like proteinase SPC6 is expressed during embryonic
RT implantation, somitogenesis, and skeletal formation.";
RL Dev. Genet. 21:75-81(1997).
CC -1- FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENOPEPTASE ACTIVITY
CC WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY. CAPABLE
CC OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF. MAY BE RESPONSIBLE
CC FOR THE MATURATION OF GASTROINTESTINAL PEPTIDES. MAY BE INVOLVED
CC IN THE CELLULAR PROLIFERATION OF ADRENAL CORTEX VIA THE ACTIVATION
CC OF GROWTH FACTORS.
CC -1- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR
CC PROPEPTIDES BY CLEAVAGE OF ARG-XAA-YAA-ARG-1-ZAA BONDS, WHERE XAA
CC CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
CC -1- SUBCELLULAR LOCATION: PCS5 IS SECRETED THROUGH THE REGULATED
CC SECRETORY PATHWAY. PCS5 IS A TYPE I MEMBRANE PROTEIN LOCALIZED TO
CC A PARANUCLEAR POST-GOLGI NETWORK COMPARTMENT IN COMMUNICATION WITH
CC EARLY ENDOSOMES.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; PCSB/LONG (SHOWN HERE)
CC AND PCS5/SHORT ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: PCS5 IS EXPRESSED IN MOST TISSUES BUT IS MOST
CC ABUNDANT IN THE INTESTINE AND ADRENALS. PCSB IS EXPRESSED IN THE
CC INTESTINE, ADRENALS AND LUNG BUT NOT IN THE BRAIN.
CC -1- DEVELOPMENTAL STAGE: WEAKLY EXPRESSED THROUGHOUT THE EMBRYO,
CC EXCEPT IN THE DEVELOPING NERVOUS SYSTEM, THE KIDNEY AND THE LIVER,
CC BUT MARKEDLY UPREGULATED AT DISCRETE SITES DURING DEVELOPMENT. AT

CC	E6.5,	PROMINENT EXPRESSION OBSERVED IN DIFFERENTIATED DECIDUA.
CC	E7.5,	INTENSE EXPRESSION IN EXTRAEMBRYONIC ENODERM, AMNION AND NASCENT MESODERM.
CC	SAC FOLLOWED BY A CONFINATION TO DERMATOGENOME COMPARTMENT.	BETWEEN E9.5 AND E11.5, ABUNDANT EXPRESSION IN AER (THICKENED ECOTDERMAL CELLS OF LIMB BUDS).
CC	TO THE CONDENSING MESCNCWM SURROUNDING THE CARTRILAGE.	AT THIS STAGE, STRONG EXPRESSION ALSO DETECTED IN VERTEBRAL AND FACIAL CARTRILAGE PRIMORDIA AND IN THE MUSCLE OF THE TONGUE.
CC	ABUNDANT EXPRESSION IN EPITHELIAL CELLS OF THE INTESTINAL VILLI.	IISOFORM A IS MOST ABUNDANT AT ALL STAGES BUT SIGNIFICANT LEVELS OF ISOFORM B OCCUR AT E12.5.
CC	- DOMAIN:	THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE ASSISTING THE FOLDING OF THE ZYMOGEN WITHIN THE ENDOPLASMIC RETICULUM.
CC	- DOMAIN:	AC 1 AND AC 2 (CLUSTERS OF ACIDIC AMINO ACIDS) CONTAIN SORTING INFORMATION. AC 1 DIRECTS TGN LOCALIZATION AND INTERACTS WITH THE TGN SORTING PROTEIN PACG-1.
CC	- SIMILARITY:	BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE SUBTILASIN FAMILY.
CC	- SIMILARITY:	CONTAINS 1 HOMO B/P DOMAIN.
CC	<hr/>	
CC	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION BETWEEN THE SWISS INSTITUTE OF BIOINFOMATICS AND THE EMBL OUTSTATION - THE EUROPEAN BIOINFOMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY IND AND FOR COMMERCIAL ENTITIES REQUIRES A LICENSE AGREEMENT (SEE http://www.isb-slb.ch/announce/or_send_an_email_to_license@lsb-slb.ch).	
CC	<hr/>	
DR	EMBL;	D17583; BA004507.1; -
DR	EMBL;	D12619; BA002143.1; -
DR	EMBL;	L14932; AAA74636.1; -
DR	PIR;	JX0248; JX0248.
DR	PIR;	A48225; A48225.
DR	HSSP;	O99405; LMPF.
DR	MEROPS;	S08.076; -
DR	MGD;	GCI:97515; PCSK5.
DR	InterPro;	IPRO00209; -
DR	InterPro;	IPRO02884; -
DR	pfam;	PF01483; P.1.
DR	pfam;	pf00082; Peptidase_S8; 1.
DR	PRINTS;	PR00723; SUBTILISIN.
DR	PROSITE;	PS00136; SUBTILASIN.ASP; 1.
DR	PROSITE;	PS00137; SUBTILASIN.HIS; 1.
DR	PROSITE;	PS00138; SUBTILASIN.SER; 1.
KW	Hydrolase on pair of basic residues; Glycoprotein; zymogen; Signal; cleavage on pair of basic residues; Repeat; Alternative splicing; Transmembrane.	
KM	SIGNAL	1
FT	PROPEP	35 116
FT	CHAIN	117 1877
FT	<hr/>	
FT	DOMAIN	117 1768
FT	TRASMEM	1769 1789
FT	DOMAIN	1790 1877
FT	DOMAIN	117 452
FT	DOMAIN	464 602
FT	DOMAIN	638 1753
FT	DOMAIN	1825 1844
FT	DOMAIN	1856 1877
FT	SITE	116 117
FT	SITE	521 523
FT	ACT_SITE	173 173
FT	ACT_SITE	214 214
FT	ACT_SITE	388 388
FT	CAROHND	227 227
FT	CAROHND	383 383
FT	CAROHND	667 667
FT	CAROHND	754 754
FT	CAROHND	804 804
FT	CAROHND	854 854
FT	CAROHND	951 951
FT	<hr/>	
FT	DOMAIN	117 1768
FT	TRANSMEM	1769 1789
FT	DOMAIN	1790 1877
FT	DOMAIN	117 452
FT	DOMAIN	464 602
FT	DOMAIN	638 1753
FT	DOMAIN	1825 1844
FT	DOMAIN	1856 1877
FT	SITE	116 117
FT	SITE	521 523
FT	ACT_SITE	173 173
FT	ACT_SITE	214 214
FT	ACT_SITE	388 388
FT	CAROHND	227 227
FT	CAROHND	383 383
FT	CAROHND	667 667
FT	CAROHND	754 754
FT	CAROHND	804 804
FT	CAROHND	854 854
FT	CAROHND	951 951
FT	<hr/>	
FT	DOMAIN	117 1768
FT	TRANSMEM	1769 1789
FT	DOMAIN	1790 1877
FT	DOMAIN	117 452
FT	DOMAIN	464 602
FT	DOMAIN	638 1753
FT	DOMAIN	1825 1844
FT	DOMAIN	1856 1877
FT	SITE	116 117
FT	SITE	521 523
FT	ACT_SITE	173 173
FT	ACT_SITE	214 214
FT	ACT_SITE	388 388
FT	CAROHND	227 227
FT	CAROHND	383 383
FT	CAROHND	667 667
FT	CAROHND	754 754
FT	CAROHND	804 804
FT	CAROHND	854 854
FT	CAROHND	951 951
FT	<hr/>	
FT	DOMAIN	117 1768
FT	TRANSMEM	1769 1789
FT	DOMAIN	1790 1877
FT	DOMAIN	117 452
FT	DOMAIN	464 602
FT	DOMAIN	638 1753
FT	DOMAIN	1825 1844
FT	DOMAIN	1856 1877
FT	SITE	116 117
FT	SITE	521 523
FT	ACT_SITE	173 173
FT	ACT_SITE	214 214
FT	ACT_SITE	388 388
FT	CAROHND	227 227
FT	CAROHND	383 383
FT	CAROHND	667 667
FT	CAROHND	754 754
FT	CAROHND	804 804
FT	CAROHND	854 854
FT	CAROHND	951

FT CARBOHYD 1016 1016 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1220 1220 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1317 1317 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1523 1523 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1711 1711 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1733 1733 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPPLIC 878 915 GEYIDDOHCONCASCACAKGPGTDEICSPYTRVLD
 (IN ISOPFORM PCSA).
 FT VARSPPLIC 916 1877 MISSING (IN ISOPFORM PCSA).
 SO SEQUENCE 1877 AA; 209287 MW; EC850E2DF20EA1C3 CRC64;

Query Match
 Best Local Similarity 75.0%; Score 36; DB 1; Length 1877;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCVOD 7
 DB 1110 GGGCVOD 1116

RESULT 9
 ID YMS4_MAIZE STANDARD; PRT; 256 AA.
 AC P10580;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE HYPOHETICAL 29 KDA PROTEIN IN MITOCHONDRIAL S-1 DNA (URF 4).
 OS Zea mays (Maize).
 OC Mitochondrion.
 OC Plasmid S-1.
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoidae;
 OC Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN SEQUENCE FROM N.A.
 RA Pallard M., Sederoff R.R., Levings C.S. III;
 RT "Nucleotide sequence of the S-1 mitochondrial DNA from the S cytoplasm
 of male.";
 RT EMBO J. 4:1125-1128(1985).
 CC -1- MISCELLANEOUS: THE MITOCHONDRIA FROM THE S MALE-STERILE CYTOPLASM
 OF MAIZE CONTAIN UNIQUE DNA-PROTEIN COMPLEXES, DESIGNATED S-1 AND
 S-2. THESE COMPLEXES CONSIST OF DOUBLE-STRANDED LINEAR DNAs WITH
 PROTEINS COVALENTLY ATTACHED TO THE 5' TERMINI.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; X02451; -; NOT_ANNOTATED.CDS.
 DR MAIZEDB: 69620.
 KW Hypothetical protein; Mitochondrion; Plasmid.
 SO SEQUENCE 256 AA; 29338 MW; BID6DB31E402D175 CRC64;

Query Match
 Best Local Similarity 72.9%; Score 35; DB 1; Length 256;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCVOD 8
 DB 214 GGGCEEG 221

RESULT 10
 ARGB_ECOLI

ID ARGB_ECOLI STANDARD; PRT; 258 AA.
 AC P11445;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ACETYLGUTAMATE KINASE (EC 2.7.2.8) (NAG KINASE) (AGK) (N-ACETYL-L-
 GLUTAMATE 5-PHOSPHOTRANSFERASE).
 GN ARGB.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=89121510; PubMed=2851495;
 RA Parsot C., Boyen A., Cohen G.N., Glandsdorff N.;
 RT "Nucleotide sequence of Escherichia coli argB and argC genes:
 comparison of N-acetylglutamate kinase and N-acetylglutamate-gamma-
 semialdehyde dehydrogenase with homologous and analogous enzymes.";
 RL Gene 68:275-283(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA SPRATN-K12 / MG1655;
 RX MEDLINE=94089392; PubMed=8265357;
 RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
 RA Daniels D.L.;
 RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
 region from 89.2 to 92.8 minutes.";
 RL Nucleic Acids Res. 21:5408-5417(1993).
 RN [3]
 RP CRYSTALLIZATION.
 RX MEDLINE=99322417; PubMed=10393305;
 RA Gil F., Ramon-Matigues S., Marina A., Pita I., Rubio V.;
 RT "N-acetyl-L-glutamate kinase from Escherichia coli: cloning of the
 gene, purification and crystallization of the recombinant enzyme and
 preliminary X-ray analysis of the free and ligand-bound forms.";
 RL Acta Crystallogr. D 55:1350-1352(1999).
 CC -1- CATALYTIC ACTIVITY: ATP + N-ACETYL-L-GLUTAMATE = ADP +
 N-ACETYL-L-GLUTAMATE 5-PHOSPHATE.
 CC -1- PATHWAY: SECOND STEP IN ARGININE BIOSYNTHESIS.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE ACETYLGUTAMATE KINASE FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; M21446; AAA23478.1; -;
 DR EMBL; U00006; AAC43065.1; -;
 DR EMBL; AE000470; AAC76941.1; -;
 DR PIR; JF0331; KIECAF.
 DR Ecogene; E310064; argB.
 DR InterPro; IPR001048; -;
 DR Pfam; PF00696; aakkinase; 1.
 KW Arginine biosynthesis; Transferase; Kinase.
 SO SEQUENCE 258 AA; 27159 MW; 8B916B8BC143738 CRC64;

Query Match
 Best Local Similarity 72.9%; Score 35; DB 1; Length 258;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCVOD 7
 DB 43 GGGCVOD 49

RESULT 11

GLGD_BACST
ID GLGD_BACST STANDARD: PRT: 343 AA.
AC 008327;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GLYCOGEN BIOSYNTHESIS PROTEIN GLGD.
GN GLGD.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
CC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1422;
[1]
RN RP
RC SEQUENCE FROM N.A.
RX STRAIN=TBRE14;
RX MEDLINE=97386405; PubMed=9244254;
RA Takata H., Takana T., Okada S., Takagi M., Imanaka T.;
RT "Characterization of a gene cluster for glycogen biosynthesis and a
heterotetrameric ADP-glucose pyrophosphorylase from Bacillus
stearothermophilus";
RT stearothermophilus";
RL J. Bacteriol. 179:4689-4698(1997).
CC -1- FUNCTION: REQUIRED FOR THE SYNTHESIS OF GLYCOGEN.
CC -1- SIMILARITY: TO ADP-GLUCOSE PYROPHOSPHORYLASE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: D87026; BAA19590.1; -
DR Glycogen biosynthesis.
KW SEQUENCE 343 AA: 38965 MW; 20A11370E10A0040 CRC64;
SQ
Query Match 70.8%; Score 34; DB 1; Length 343;
Best Local Similarity 75.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 GGCVCVDG 8
DB 299 GGCVCVDG 306
| | | | |
| | | | |
RESULT 12
CIRL_CITFR STANDARD: PRT: 618 AA.
AC P23182;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-NOV-1991 (Rel. 20, Last annotation update)
DE CITROLYSIN PROTEIN 1.
OS Citrobacter freundii.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Citrobacter.
OX NCBI_TaxID=546;
[1]
RN RP
RC SEQUENCE FROM N.A.
RX al Zaag A., Pemberton J.M.;
RT "DNA sequence of haemolysin gene from Citrobacter freundii.";
RL FEBS Microbiol. Lett. 49:131-135(1988).
CC -1- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CC CELL MEMBRANS AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
CC DEFINED.
DR PIR: S06446; S06446.
DR Hemolysis.
KW SEQUENCE 618 AA: 70471 MW; E13D17CA42FCD963 CRC64;
SQ
Query Match 70.8%; Score 34; DB 1; Length 618;
Best Local Similarity 71.4%; Pred. No. 85;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 GGCVCVD 7
DB 287 GGCCLDD 293
| | | | |
| | | | |
RESULT 13
YC92_CAEEL STANDARD: PRT: 621 AA.
AC P55115;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL ZINC METALLOPROTEINASE T04G9.2 (EC 3.4.24.-).
GN T04G9.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
CC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN RP
RC SEQUENCE FROM N.A.
RX STRAIN=BRISTOL N2;
RA Minx P.;
RL Submitted (NOV-1995) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A (ZINC
CC METALLOPROTEASE); ALSO KNOWN AS THE ASTACIN SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: U41274; AAA82461.1; -
DR HSSP: P07584; 1IAD.
DR WormPep: T04G9.2; CE04883.
DR InterPro: IPR00130; -
DR InterPro: IPR001506; -
DR Pfam: PF01400; Astacin; 1.
DR PRINTS: PR00480; ASTACIN.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Hypothetical protein; Hydrolase; Metalloprotease; Zinc.
FT METAL 255 255 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 256 256 BY SIMILARITY.
FT METAL 259 259 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 265 265 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 621 AA: 68584 MW; A98D3721C8ADE489 CRC64;
OY 1 GGCVCVDG 8
DB 244 GSGCIOKG 251
| | | | |
| | | | |
RESULT 14
TKTL_KLULA STANDARD: PRT: 679 AA.
AC Q12630;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE TRANSKETOLASE (EC 2.2.1.1) (TK).
GN TKTL.
OS Kluyveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.

OX NCBI_TaxID=28985;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CBS 2359/152;
 RX MEDLINE=97153463; PubMed=9000376;
 RA Jacoby J.J., Helms J.J.;
 RT "Analysis of a transketolase gene from Kluyveromyces fragilis reveals
 RL that the yeast enzymes are more related to transketolases of
 CC Prokaryotic origins than to those of higher eukaryotes";
 CC Curr. Genet. 31:15-21(1997).
 CC -1- CATALYTIC ACTIVITY: SEDOHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE
 CC -1- COFACTOR: THIAMINE PYROPHOSPHATE.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U65983; AAB05935.1; -
 DR HSSP: P23254; IAYU.
 DR InterPro: IPR000360; -
 DR Pfam: PF00456; transketolase; 1.
 DR PROSITE: PS00801; TRANSKETOLASE_1; 1.
 DR PROSITE: PS00802; TRANSKETOLASE_2; 1.
 DR Transferrase: Thiamine pyrophosphate.
 KW SEQUENCE 679 AA; 73703 MW; BCSF3CF61A78CDA4 CRC64;
 SQ
 Query Match 70.8%; Score 34; DB 1; Length 679;
 Best Local Similarity 62.5%; Pred. No. 92;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 GGGCVDG 8
 | | | | |
 Db 156 GDCGDEG 163
 RESULT 15
 ID TKT1_YEAST STANDARD: PRT; 679 AA.
 AC P23254;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE TRANSKETOLASE 1 (EC 2.2.1.1) (TK 1).
 GN TKT1 OR YPR074C OR YP9499.29C.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 [1]
 RN SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RP MEDLINE=92144611; PubMed=1737042;
 RA Fletcher T.S., Kwee I.L., Nakada T., Laigman C., Martin B.M.;
 RT "DNA sequence of the yeast transketolase gene";
 RL Biochemistry 31:1892-1896(1992).
 CC [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=M303-1A;
 RX MEDLINE=94043273; PubMed=8226984;
 RA "Yeast TKT1 gene encodes a transketolase that is required for
 RT efficient glycolysis and biosynthesis of aromatic amino acids";
 RL J. Biol. Chem. 268:24346-24352(1993).
 CC [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;

RA Badcock K., Churcher C.M., Barrell B.G., Rajandream M.A.;
 RL Submitted (May-1995) to the EMBL/GenBank/DBJ databases.
 [4]
 RN SEQUENCE FROM N.A.
 RP STRAIN=S288C / AB972;
 RC STRAIN=S288C / AB972;
 RA Johnston A., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
 RA Favellon A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,
 RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,
 RA Johnston L., Langston Y., Latreille P., Le T., Mardis E., Menezes S.,
 RA Miller N., Nhan M., Pauley A., Peluso D., Rinken L., Riles L.,
 RA Talbot A., Trevisan E., Vignati D., Wilcox L., Woldman P., Vaudin M.,
 RA Wilson R., Waterson R.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 1-35.
 RX MEDLINE=92253546; PubMed=1812485;
 RA Nixon P.F., Dugdaley R.G.;
 RT "The N-terminal amino acid sequence of yeast transketolase";
 RL Protein Seq. Data Anal. 4:325-326(1991).
 RN [6]
 RX X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RA Lindqvist Y., Schneider G., Emmler U., Sundstroem M.;
 RT "Three-dimensional structure of transketolase, a thiamine diphosphate
 RL dependent enzyme, at 2.5-A resolution";
 RN EMBL: J. 11:2373-2379(1992).
 [7]
 RX X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RA MEDLINE=94231561; PubMed=8176731;
 RN Nikkola M., Lindqvist Y., Schneider G.;
 RT "Refined structure of transketolase from Saccharomyces cerevisiae at
 RL 2.0-A resolution";
 RN J. Mol. Biol. 238:387-404(1994).
 [8]
 RP MUTAGENESIS OF HIS-102.
 RX MEDLINE=96062276; PubMed=9398292;
 RA Wikner C., Meshalkina L., Nilsson U., Baackstroem S., Lindqvist Y.,
 RA Schneider G.;
 RT "Identification of catalytically important residues in yeast
 RL transketolase";
 RN Biochemistry 36:15643-15649(1997).
 [9]
 RP MUTAGENESIS OF HIS-102.
 RX MEDLINE=96085137; PubMed=8521838;
 RA Wikner C., Meshalkina L., Nilsson U., Baackstroem S., Lindqvist Y.,
 RA Schneider G.;
 RT "His103 in yeast transketolase is required for substrate recognition
 RL and catalysis";
 RN Eur. J. Biochem. 233:750-755(1995).
 CC -1- CATALYTIC ACTIVITY: SEDOHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE
 CC -1- COFACTOR: THIAMINE PYROPHOSPHATE. REQUIRES MAGNESION IONS FOR
 CC CATALYTIC ACTIVITY.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M63302; AAA35168.1; -
 DR EMBL: X73324; CAAS1693.1; -
 DR EMBL: 249219; CAAB9191.1; -
 DR EMBL: 271255; CAAB94982.1; -
 DR EMBL: U51033; AAB68125.1; -
 DR PIR: A42084; XUBYTK.
 DR PIR: S21067; S21067.
 DR PDB: ITRK; 3O-APR-94.
 DR PDB: ITRK; 3O-NOV-94.

DR PDB: 1TRK; 30-NOV-94.
 DR PDB: 1TRK; 30-NOV-94.
 DR PDB: 1NGS; 12-FEB-97.
 DR PDB: 1AY0; 13-MAY-98.
 DR SWISS-2DPAGE; P23254; YEAST.
 DR YEPD; 4776; -
 DR SGD; S0006278; TKL1.
 DR InterPro: IPR000360; -
 DR pfam: PF00456; transketolase; 1.
 DR PROSITE: PS00801; TRANSKETOLASE_1; 1.
 DR PROSITE: PS00802; TRANSKETOLASE_2; 1.
 KW Transferrase; Thiamine pyrophosphate; Magnesium; Multigene family;
 3D-structure.
 FT INIT.MET 0 0
 FT ACT_SITE 102 102 PROBABLE.
 FT MUTAGEN 102 102 H->A,F,N; LOSS OF ACTIVITY.
 FT CONFLICT 36 37 MA->RS (IN REF. 1).
 FT CONFLICT 44 76 WQMRNPNTPDMINDREVLNSGHVALLYSM -> GSEN
 (IN REF. 1).
 FT CONFLICT 44 76 AHEPNOPKGTSEIDLSCLTVTRSLCCY
 (IN REF. 1).
 FT CONFLICT 135 142 AATYKRG -> DMPETTSRA (IN REF. 1).
 FT CONFLICT 231 233 AQA -> RQR (IN REF. 1).
 FT CONFLICT 242 256 LKMTTITIGYSLHA -> FDQNDHNHWLRLRS
 (IN REF. 1).
 FT CONFLICT 382 382 L -> LVLPIL (IN REF. 1).
 FT CONFLICT 395 395 D -> S (IN REF. 1).
 FT CONFLICT 527 537 RQNLPOLEBS -> PDKTCHMKVAL (IN REF. 1).
 FT CONFLICT 527 537 SGRAPVEVEFEFTPEGVAERAKTIAFYKGLISPLKA
 F -> PVRHOKSSSSSVSPQKVLKELKRLPHSHIRVTS
 (IN REF. 1).
 FT CONFLICT 638 679
 FT
 FT
 SO SEQUENCE 679 AA; 73674 MW; E5A148CD604C937D CRC64;

Query Match
 Best Local Similarity 70.8%; Score 34; DB 1; Length 679;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCVQDG 8
 DB 155 GDGCIQEG 162

Search completed: June 13, 2001, 14:30:36
 Job time: 526 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:29:40 ; Search time 150.43 Seconds
(without alignments)
6.233 Million cell updates/sec

Title: PCT-US01-05825A-22

Sequence: 1 GCGCVDG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_PROTOZOA:*
12: SP_UNCLASSIFIED:*
13: SP_VIRUS:*
14: SP_VIRUS:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	87.5	206	13 090802	090802 gallus galli
2	39	81.2	196	2 084353	084353 chlamydia t
3	39	81.2	196	2 09PK45	09PK45 chlamydia m
4	38	79.2	250	10 09LG70	09LG70 oryza sativ
5	38	79.2	302	10 041801	041801 zea mays (m
6	38	79.2	653	5 09VZK2	09VZK2 drosophila
7	38	79.2	664	5 024140	024140 drosophila
8	38	79.2	2395	5 027167	027167 parametium
9	37	77.1	110	14 067656	067656 garlic late
10	37	77.1	110	14 067553	067553 garlic late
11	37	77.1	279	1 027091	027091 methanobact
12	37	77.1	379	5 09U5K8	09U5K8 schistocerc
13	37	77.1	2717	5 094710	094710 parametium
14	37	77.1	2721	5 076973	076973 parametium
15	36	75.0	103	4 099227	099227 homo sapien
16	36	75.0	109	5 09W2B6	09W2B6 drosophila
17	36	75.0	119	6 077753	077753 canis famill
18	36	75.0	145	1 026242	026242 methanobact
19	36	75.0	161	5 09VZ39	09VZ39 drosophila

20	36	75.0	352	2 050170	050170 mycobacteri
21	36	75.0	362	2 006408	006408 mycobacteri
22	36	75.0	434	5 017739	017739 caenorhabd
23	36	75.0	452	2 09KE17	09KE17 bacillus ha
24	36	75.0	583	10 022889	022889 arabidopsis
25	36	75.0	640	14 011308	011308 molluscum c
26	36	75.0	680	14 098187	098187 molluscum c
27	36	75.0	1081	5 09W3R9	09W3R9 drosophila
28	36	75.0	1193	5 019617	019617 caenorhabd
29	36	75.0	1267	5 09U0U5	09U0U5 leishmania
30	36	75.0	1487	4 014047	014047 homo sapien
31	36	75.0	1491	13 091718	091718 xenopus lae
32	36	75.0	1548	11 062040	062040 mus musculu
33	36	75.0	2233	5 094711	094711 parametium
34	36	75.0	2248	5 09VY17	09VY17 drosophila
35	35	72.9	231	1 026548	026548 methanobact
36	35	72.9	256	8 09MJC3	09MJC3 zea mays (m
37	35	72.9	273	4 09UHF1	09UHF1 homo sapien
38	35	72.9	622	5 09VHM5	09VHM5 drosophila
39	35	72.9	1218	5 09VHR6	09VHR6 drosophila
40	35	72.9	1722	5 019350	019350 caenorhabd
41	34.5	71.9	165	4 09NR65	09NR65 homo sapien
42	34.5	71.9	748	4 09NR64	09NR64 homo sapien
43	34.5	71.9	749	4 09P238	09P238 homo sapien
44	34	70.8	296	2 050407	050407 mycobacteri
45	34	70.8	447	2 053137	053137 rhodococcus

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	206 AA.
ID	090802			
AC	090802			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)			
DE	COLLAGEN TYPE 2 PRECURSOR (FRAGMENT).			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92078225; PubMed=1744138;			
RA	Nah H.D., Upholt W.B.;			
RT	"Type II collagen mRNA containing an alternatively spliced exon predominates in the chick limb prior to chondrogenesis.";			
RL	J. Biol. Chem. 266:23446-23452(1991).			
DR	EMBL; M74435; AAA48714.1; -			
DR	INTERPRO; IPR000087; -			
DR	INTERPRO; IPR001007; -			
DR	PFAM; PF01093; vwc; 1.			
DR	PFAM; PF01391; Collagen; 1.			
DR	PROSITE; PS01208; vwc; 1.			
KW	SIGNAL.			
FT	CHAIN 1 25			POTENTIAL.
FT	CHAIN 185 >206			COLLAGEN TYPE 2.
FT	NON_TER 206 206			
SQ	SEQUENCE 206 AA; 20793 MW; E3EB84ACF0C31459 CRC64;			

Query Match 87.5%; Score 42; DB 13; Length 206;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GCGVDG 8
Db 35 GCGVDG 41

```

RESULT 2
*ID 084353 PRELIMINARY; PRT: 196 AA.
DC 084353;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DE 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
GN MAF-TYPE PROTEIN.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UM-3/CX;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis."
RL Science 282:754-759(1998).
DR EMBL: AE001308; AAC67944.1;
SQ SEQUENCE 196 AA; 21978 MW; EDB6CFD52F93073 CRC64;

Query Match
Best Local Similarity 81.2%; Score 39; DB 2; Length 196;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCVCVODG 8
DB 150 GGCVCVODG 157

RESULT 3
*ID 09PK45 PRELIMINARY; PRT: 196 AA.
DC 09PK45;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
GN MAF PROTEIN.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WOPN / NIGG;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Ungerback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=H84; TISSUE=LEAF;
RX MEDLINE=96032831; PubMed=7567449;
RA Dodson R., Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G.,
RA Salzberg S.L., Eisen J., Fraser C.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE002331; AAF39457.1;
RT TIGR: TC0628;
SQ SEQUENCE 196 AA; 21730 MW; 9886B059E52D78A5 CRC64;

Query Match
81.2%; Score 39; DB 2; Length 196;

```

```

RESULT 4
*ID 09LGTO PRELIMINARY; PRT: 250 AA.
DC 09LGTO;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
GN P0489A01.11 PROTEIN.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RL clone:P0489A01."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF002484; BAA99520.1;
SQ SEQUENCE 250 AA; 28494 MW; 27574C02BA14BCF3 CRC64;

Query Match
Best Local Similarity 79.2%; Score 38; DB 10; Length 250;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCVCVODG 8
DB 17 GAGCVKDG 24

RESULT 5
*ID 041801 PRELIMINARY; PRT: 302 AA.
DC 041801;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-NOV-1996 (TREMBlrel. 12, Last annotation update)
GN DOP3 GENE (FRAGMENT).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H84; TISSUE=LEAF;
RA Shuichi Y.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=H84; TISSUE=LEAF;
RX MEDLINE=96032831; PubMed=7567449;
RA Yanagisawa S.;
RT "A novel DNA-binding domain that may form a single zinc finger
RT motif."
RL Nucleic Acids Res. 23:3403-3410(1995).
DR EMBL: X79935; CAA56288.1;
DR MENDEL: 11281; Zeama;1274;11281.
RT NON_TER
SQ SEQUENCE 302 AA; 31197 MW; B7826F297211B9AD CRC64;

Query Match
79.2%; Score 38; DB 10; Length 302;

```


Best Local Similarity 75.0%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGCVODG 8
Db 240 GGGCGEDG 247

RESULT 6
O9VZK2 PRELIMINARY; PRT; 653 AA.
AC O9VZK2: O9Y144;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE SCRT PROTEIN (SCRATCH).
GN SCRT.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amathides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang O., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
RA Abril J.F., Abmayr A., An H.-J., Andrews-Plamkoc C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borliva D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dev I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foaier C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harlow N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Liao P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mlshina N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Put V., Reese M.G.,
RA Relneit K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spredling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Welschbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT *The genome sequence of Drosophila melanogaster.*;
RL Science 287:2185-2195(2000).
RN [12]
RP SEQUENCE OF 1-567 FROM N.A.
RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Broksstein P., Tsang G.,
RA Agbayani A., Arcaina T.T., Baxter E., Blazed R.G., Butenoff C.,
RA Chame M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
RA Calle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm J.,
RA Houston K.A., Humast S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,
RA Park S., Sequeira A., Sethi H., Snir E., Svitskas R.R., Weinburg T.,
RA Celniker S.E.;

RT "Full length Drosophila melanogaster cDNA sequence."
RL Submitted (Apr-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AE003479; AAF47819.1; -
DR EMBL: AF145627; AAD38602.1; -
DR HSSP: P25490; 1UBD.
DR FLYBASE: FBgn0004880; scrt.
DR INTERPRO: IPR000822; -
DR PFAM: PF00096; zf-C2H2; 5.
DR PROSITE: PS00028; ZINC_FINGER_C2H2; 4.
KW Zinc-finger; Metal-binding; DNA-binding.
SQ SEQUENCE 653 AA; 70774 MW; 15C27EE93F3F45D CRC64;

Query Match 79.2%; Score 38; DB 5; Length 653;
Best Local Similarity 75.0%; Pred. No. 49;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 GGGCVODG 8
Db 408 GGGCLODG 415

RESULT 7
O24140 PRELIMINARY; PRT; 664 AA.
AC O24140;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE NEURON SPECIFIC ZINC FINGER TRANSCRIPTION FACTOR.
GN SCRATCH.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96010286; PubMed=7557390;
RA Roark M., Sturtevant M.A., Emery J., Vaessin H., Grell E., Bier E.;
RT "Scratch, a pan-neural gene encoding a zinc finger protein related to
RT snail, promotes neuronal development."
RL Genes Dev. 9:2384-2398(1995).
DR EMBL: U36477; AAA91035.1; -
DR HSSP: P25490; 1UBD.
DR FLYBASE: FBgn0004880; scrt.
DR INTERPRO: IPR000822; -
DR PFAM: PF00096; zf-C2H2; 5.
DR PRINTS: PR00048; ZINC_FINGER.
DR PROSITE: PS00028; ZINC_FINGER_C2H2; 4.
KW Zinc-finger; Metal-binding; DNA-binding.
SQ SEQUENCE 664 AA; 72569 MW; 687F2DAA8FE8968F CRC64;

Query Match 79.2%; Score 38; DB 5; Length 664;
Best Local Similarity 75.0%; Pred. No. 50;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 GGGCVODG 8
Db 408 GGGCLODG 415

RESULT 8
O27167 PRELIMINARY; PRT; 2395 AA.
AC O27167;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE 51B TYPE SURFACE PROTEIN.
OS Paramoecium tetraurelia.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Penicillida;

OC Parametrium
 OX NCBI_TaxID=5888;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-STOCK 51;
 RX MEDLINE=9509630; PubMed=7800503;
 RA Scott J., Leack C., Forney J.;
 RT "Analysis of the micronuclear B type surface protein gene in
 RL Parametrium tetraurelia.";
 RN Nucleic Acids Res. 22:5079-5084(1994).
 RP [2]
 RC SEQUENCE FROM N.A.
 RN STRAIN-STOCK 51;
 RA Scott J.M., Leack C.L., Forney J.D.;
 RL Genetics 133:189-198(1993).
 DR EMBL: U07603; AAA81947.1;
 DR EMBL: L04795; AAA16710.1;
 DR INTERPRO: IPR000950;
 DR INTERPRO: IPR001680;
 DR INTERPRO: IPR002895;
 DR PFAM: PF01508; Parametrium_SA; 28.
 DR PROSITE: PS00241; RECEPTOR_CYTOKINES_1; UNKNOWN_1.
 DR PROSITE: PS00678; WD_REPEATS; UNKNOWN_1.
 SO SEQUENCE 2395 AA; 246430 MW; 8633E022649B2D61 CRC64;

Query Match
 Best Local Similarity 79.2%; Score 38; DB 5; Length 2395;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGCYODG 8
 DB 327 GGCYVNG 334

RESULT 9
 ID 067656 PRELIMINARY; PRT; 110 AA.
 AC 067656;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE 12 KDa PROTEIN.
 OS garlic latent virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Carlavirus.
 OX NCBI_TaxID=12458;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MONJU;
 RA Jin Nam C., Sang Ik S., Jong Tae S., Jong Seob L., Yang Do C.;
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z68502; CAAG2817.1;
 DR INTERPRO: IPR001896;
 DR PFAM: PF01307; Plant_vir_prot; 1.
 DR PRODOM: PD001561; -; 1.
 SO SEQUENCE 110 AA; 11831 MW; 0BBE1C116749502C CRC64;

Query Match
 Best Local Similarity 77.1%; Score 37; DB 14; Length 110;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GGCYODG 8
 DB 47 GGCYODG 53

RESULT 10
 ID 067553 PRELIMINARY; PRT; 110 AA.
 AC 067553;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE ORF.
 OS garlic latent virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Carlavirus.
 OX NCBI_TaxID=12458;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tsuneyoshi T., Suni S.;
 RT "Differentiation among garlic viruses in mixed infections based on RT-PCR procedures and direct tissue blotting immunoassays.";
 RL Phytopathology 86:253-259(1996).
 DR EMBL: D1161; BAA01933.1;
 DR INTERPRO: IPR001896;
 DR PFAM: PF01307; Plant_vir_prot; 1.
 DR PRODOM: PD001561; -; 1.
 SO SEQUENCE 110 AA; 11917 MW; D90F69ABBF7F1CC3 CRC64;

Query Match
 Best Local Similarity 77.1%; Score 37; DB 14; Length 110;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GGCYODG 8
 DB 47 GGCYODG 53

RESULT 11
 ID 027091 PRELIMINARY; PRT; 279 AA.
 AC 027091;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
 DE ENDONUCLEASE IV.
 GN MTH1010.
 OS Methanobacterium thermoautotrophicum
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
 OC Methanobacterium.
 OX NCBI_TaxID=2166;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DELTA H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jivani N., Carnes A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RL J. Bacteriol. 179:7135-7155(1997).
 DR EMBL: AE000874; AAB85506.1;
 DR INTERPRO: IPR001719;
 DR PFAM: PF01261; AP_endonuclease2; 1.
 DR KW Endonuclease.
 SO SEQUENCE 279 AA; 31032 MW; 33C0D8572E961A83 CRC64;

Query Match
 Best Local Similarity 77.1%; Score 37; DB 1; Length 279;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCYODG 7
 DB 185 GGCYODG 191

RESULT 12
 ID 0905K8 PRELIMINARY; PRT; 379 AA.
 ID 0905K8

AC 0905K8:
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE GROUP 3 HOMEDOMAIN PROTEIN.
 GN HOX3.
 OS Schistocerca gregaria (Desert locust).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caellifera;
 CC Acridomorpha; Acridoidea; Acrididae; Schistocerca.
 OX NCBI_Taxid=7010;
 GN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CAMBRIDGE: TISSUE-EMBRYO;
 RA Falciandl F., Hausdorf B., Schroeder R., Tautz D., Akam M., Denell R.,
 RA Brown S.;
 RT "Class 3 genes in insects and the origin of zen."
 RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: X92654; CAB61208.1; -
 DR HSSP: P02833; ISAN.
 DR INTERPRO: IPR000047; -
 DR INTERPRO: IPR001356; -
 DR PFAM: PF00046; homeobox.1.
 DR PRINTS: PRO0024; HOMEBOX.
 DR PRINTS: PRO0031; HTREPRESSR.
 DR PROSITE: PS00027; HOMEBOX_1; 1.
 DR PROSITE: PS00771; HOMEBOX_2; 1.
 KW Homeobox; DNA-binding; Nuclear protein.
 SQ SEQUENCE 379 AA; 41448 MW; 95950C5B494D19A8 CRC64;

Query Match 77.1%; Score 37; DB 5; Length 379;
 Best Local Similarity 75.0%; Pred. No. 43;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGGCVQDG 8
 |||||
 Db 190 GGGCRLDG 197

RESULT 13
 Q94710 PRELIMINARY; PRT; 2717 AA.
 ID 094710:
 AC 094710:
 DT 01-FEB-1997 (Tremblrel. 02, Created)
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
 DE 51A SURFACE PROTEIN.
 OS Paramacium tetraurelia.
 CC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
 CC Paramacium.
 OX NCBI_Taxid=5888;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92106337; PubMed=1762150;
 RA Nielsen E., You Y., Forney J.;
 RT "Cysteine residue periodicity is a conserved structural feature of
 RT variable surface proteins from Paramacium tetraurelia.";
 RL J. Mol. Biol. 222:835-841(1991).
 DR EMBL: M65163; AAA61739.1; -
 DR INTERPRO: IPR001680; -
 DR INTERPRO: IPR002895; -
 DR PFAM: PF01508; Paramacium_SA; 33.
 DR PROSITE: PS00678; WD_REPEATS; UNKNOWN_1.
 SQ SEQUENCE 2717 AA; 280014 MW; 0A271E270868CAA9 CRC64;

Query Match 77.1%; Score 37; DB 5; Length 2717;
 Best Local Similarity 75.0%; Pred. No. 3e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 1 GGGCVQDG 8
 |||||

Db 195 GGGCVDSG 202

RESULT 14
 ID 076973 PRELIMINARY; PRT; 2721 AA.
 AC 076973:
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
 DE SURFACE ANTIGEN G.
 GN 51G.
 OS Paramacium tetraurelia.
 CC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
 CC Paramacium.
 OX NCBI_Taxid=5888;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-51;
 RA Dubarcourt S., Keller A.M., Meyer E.;
 RT "Homology-dependent maternal inhibition of developmental excision of
 RT internal Eliminated Sequences in Paramacium tetraurelia.";
 RL Submitted (AUG-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AJ010441; CAA09180.1; -
 DR INTERPRO: IPR000719; -
 DR INTERPRO: IPR001680; -
 DR INTERPRO: IPR002895; -
 DR PFAM: PF01508; Paramacium_SA; 33.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR PROSITE: PS00678; WD_REPEATS; UNKNOWN_1.
 SQ SEQUENCE 2721 AA; 280341 MW; 630780EB7B76242 CRC64;

Query Match 77.1%; Score 37; DB 5; Length 2721;
 Best Local Similarity 75.0%; Pred. No. 3.1e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGGCVQDG 8
 |||||
 Db 194 GGGCVDSG 201

RESULT 15
 Q99227 PRELIMINARY; PRT; 103 AA.
 ID 099227:
 AC 099227:
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
 DE COLLAGEN ALPHA 1(I) CHAIN PRECURSOR (FRAGMENT).
 GN COL2A1.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91184811; PubMed=2081599;
 RA Ryan M.C., Sieraski M., Sandell L.J.;
 RT "The human type II procollagen gene: identification of an additional
 RT protein-coding domain and location of potential regulatory sequences
 RT in the promoter and first intron.";
 RL Genomics 8:41-48(1990).
 DR [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92321501; PubMed=1671990;
 RA Adams M.E., Huang D.O., Yao L.T.,
 RT "Extraction and isolation of mRNA from adult articular cartilage.";
 RL Anal. Biochem. 202:89-95(1992).
 DR EMBL: M60299; AAA73873.1; -
 DR INTERPRO: IPR001007; -
 DR PFAM: PF00093; vwc.1.
 DR PROSITE: PS01208; WVEC; 1.

KW Signal; Glycoprotein; Cartilage.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 103 COLLAGEN ALPHA 1(II) CHAIN.
FT NON_TER 103 103
SQ SEQUENCE 103 AA; 11039 MW; E5E8CD2167086AB CRC64;

Query Match 75.0%; Score 36; DB 4; Length 103;
Best Local Similarity 85.7%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GGCVCQDG 8
1 11111
Db 32 GSCVCQDG 38

Search completed: June 13, 2001, 14:29:41
Job time: 542 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:27:06 ; Search time 78.71 Seconds
(without alignments)
1.953 Million cell updates/sec

Title: PCT-US01-05825A-22

Perfect score: 48

Sequence: 1 GGGCVQDG 8

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCtUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	75.0	493	1	US-07-615-448A-7
2	36	75.0	493	1	US-08-196-361-7
3	36	75.0	493	2	US-08-446-934-7
4	36	75.0	493	2	US-08-448-128-7
5	36	75.0	493	2	US-08-948-703-7
6	34	70.8	599	2	US-08-426-125-1
7	34	70.8	599	2	US-08-455-355-1
8	34	70.8	648	1	US-08-321-356-2
9	34	70.8	648	2	US-08-321-356-4
10	33	68.8	543	2	US-08-469-412A-7
11	33	68.8	543	4	US-09-021-715-7
12	32	66.7	14	1	US-08-143-365A-3
13	32	66.7	19	1	US-08-143-365A-10
14	32	66.7	119	1	US-08-318-157B-2
15	32	66.7	201	1	US-08-418-570-2
16	32	66.7	282	3	US-08-911-853-37
17	32	66.7	299	2	US-08-701-191A-13
18	32	66.7	300	2	US-08-701-191A-31
19	32	66.7	310	2	US-08-701-191A-6
20	32	66.7	310	2	US-08-701-191A-17
21	32	66.7	313	1	US-08-278-089A-1
22	32	66.7	313	2	US-08-838-957A-16
23	32	66.7	315	2	US-08-701-191A-2
24	32	66.7	316	4	US-07-791-931-8
25	32	66.7	337	6	5202251-1
26	32	66.7	351	2	US-08-701-191A-3
27	32	66.7			Sequence 3, Appli

28	32	66.7	378	1	US-08-070-165F-8	Sequence 8, Appli
29	32	66.7	378	2	US-08-885-418-8	Sequence 8, Appli
30	32	66.7	388	1	US-08-070-165F-4	Sequence 4, Appli
31	32	66.7	388	2	US-08-885-418-4	Sequence 4, Appli
32	32	66.7	501	2	US-08-660-963-13	Sequence 13, Appli
33	32	66.7	526	1	US-08-471-570-4	Sequence 4, Appli
34	32	66.7	599	2	US-08-426-125-3	Sequence 3, Appli
35	32	66.7	599	2	US-08-455-355-3	Sequence 3, Appli
36	32	66.7	638	2	US-08-426-125-4	Sequence 4, Appli
37	32	66.7	638	2	US-08-426-125-5	Sequence 5, Appli
38	32	66.7	638	2	US-08-455-355-4	Sequence 4, Appli
39	32	66.7	638	2	US-08-455-355-5	Sequence 5, Appli
40	32	66.7	643	1	US-08-471-570-6	Sequence 6, Appli
41	32	66.7	652	1	US-08-471-570-10	Sequence 10, Appli
42	32	66.7	729	1	US-07-640-029-3	Sequence 3, Appli
43	32	66.7	729	1	US-08-070-165F-6	Sequence 6, Appli
44	32	66.7	729	2	US-08-885-418-6	Sequence 6, Appli
45	32	66.7	731	1	US-07-921-807B-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-07-615-448A-7
Sequence 7, Application US/07615448A
Patent No. 5347076
GENERAL INFORMATION:
APPLICANT: Pohlentz, Hans-Dieter
APPLICANT: Boldol, Werner
APPLICANT: Streiber, Wolfgang
TITLE OF INVENTION: PROCESS FOR THE ISOLATION AND
CHARACTERISATION OF A GENE ENZYME SYSTEM FOR INACTIVATION
OF THE HERBICIDE PHENMEDIPHAN AND TRANSFER OF THE GENE
TITLE OF INVENTION: INO PLANTS TO PRODUCE HERBICIDE-TOLERANT PLANTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ostrolenk, Faber, Gerb & Sofien
STREET: 1180 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-8403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/615,448A
FILING DATE: 19901119
CLASSIFICATION: 800
PRIOR APPLICATION NUMBER: DE P3817384.0
APPLICATION NUMBER: DE P3817384.0
FILING DATE: 19-APR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/353,871
FILING DATE: 18-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: Dennis, Manette
REGISTRATION NUMBER: 30,623
REFERENCE/DOCKET NUMBER: F-8366(CIP)(951-64)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-382-0888
TELEFAX: 212-382-0888
TELEX: 236925
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 493 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-615-448A-7

Query Match 75.0%; Score 36; DB 1; Length 493;
Best Local Similarity 71.4%; Pred. No. 2.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGCYVDG 8
111111
DB 15 GGCLEDDG 21

RESULT 2
US-08-196-361-7
; Sequence 7, Application US/08196361
; Patent No. 5543306
; GENERAL INFORMATION:
; APPLICANT: Pohlentz, Hans-Dieter
; APPLICANT: Boldol, Werner
; APPLICANT: Streber, Wolfgang
; TITLE OF INVENTION: PROCESS FOR THE ISOLATION AND
; TITLE OF INVENTION: CHARACTERISATION OF A GENE ENZYME SYSTEM FOR INACTIVATION
; TITLE OF INVENTION: OF THE HERBICIDE PHENMEDIPHAN AND TRANSFER OF THE GENE
; TITLE OF INVENTION: INTO PLANTS TO PRODUCE HERBICIDE-TOLERANT PLANTS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
; STREET: 1180 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-8403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/196,361
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P3817384.0
; FILING DATE: 19-MAY-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/353,871
; FILING DATE: 18-MAY-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/615,448
; FILING DATE: 19-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Dennis, Manette
; REGISTRATION NUMBER: 30,623
; REFERENCE/DOCKET NUMBER: FA-1623 DIV (951-87)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-382-0700
; TELEFAX: 212-382-0888
; TELEX: 236925
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 493 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-196-361-7

Query Match 75.0%; Score 36; DB 1; Length 493;
Best Local Similarity 71.4%; Pred. No. 2.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGCYVDG 8
111111
DB 15 GGCLEDDG 21

RESULT 3
US-08-446-934-7
; Sequence 7, Application US/08446934
; Patent No. 5846803
; GENERAL INFORMATION:
; APPLICANT: Pohlentz, Hans-Dieter
; APPLICANT: Boldol, Werner
; APPLICANT: Streber, Wolfgang
; TITLE OF INVENTION: PROCESS FOR THE ISOLATION AND
; TITLE OF INVENTION: CHARACTERISATION OF A GENE ENZYME SYSTEM FOR INACTIVATION
; TITLE OF INVENTION: OF THE HERBICIDE PHENMEDIPHAN AND TRANSFER OF THE GENE
; TITLE OF INVENTION: INTO PLANTS TO PRODUCE HERBICIDE-TOLERANT PLANTS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
; STREET: 1180 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-8403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,934
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/196,361
; FILING DATE: 02-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/615,448
; FILING DATE: 19-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/353,871
; FILING DATE: 18-MAY-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P3817384.0
; FILING DATE: 19-MAY-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Meliman, Edward A.
; REGISTRATION NUMBER: 24,735
; REFERENCE/DOCKET NUMBER: P/951-114
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-382-0700
; TELEFAX: 212-382-0888
; TELEX: 236925
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 493 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-446-934-7

Query Match 75.0%; Score 36; DB 2; Length 493;
Best Local Similarity 71.4%; Pred. No. 2.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGCYVDG 8
111111
DB 15 GGCLEDDG 21

RESULT 4
US-08-448-128-7
; Sequence 7, Application US/08448128
; Patent No. 5948670

GENERAL INFORMATION:
APPLICANT: Pohlenz, Hans-Dieter
APPLICANT: Boisdol, Werner
APPLICANT: Stredler, Wolfgang
TITLE OF INVENTION: PROCESS FOR THE ISOLATION AND
TITLE OF INVENTION: CHARACTERISATION OF A GENE ENZYME SYSTEM FOR INACTIVATION
TITLE OF INVENTION: OF THE HERBICIDE PHENMEDIPHAN AND TRANSFER OF THE GENE
TITLE OF INVENTION: INTO PLANTS TO PRODUCE HERBICIDE-TOLERANT PLANTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
STREET: 1180 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-8403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448.128
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/196.361
FILING DATE: 02-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/615.448
FILING DATE: 19-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/353.871
FILING DATE: 18-MAY-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P3817384.0
FILING DATE: 19-MAY-1988
ATTORNEY/AGENT INFORMATION:
NAME: Meliman, Edward A.
REGISTRATION NUMBER: 24,735
REFERENCE/DOCKET NUMBER: P/951-113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-382-0700
TELEFAX: 212-382-0888
TELEX: 236925
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 493 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-448-128-7

Query Match 75.0%; Score 36; DB 2; Length 493;
Best Local Similarity 71.4%; Pred. No. 2.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGCVDG 8
DB 15 GGCLEDG 21

RESULT 5
US-08-948-703-7
Sequence 7, Application US/08948703
Patent No. 6060645
GENERAL INFORMATION:
APPLICANT: Pohlenz, Hans-Dieter
APPLICANT: Boisdol, Werner
APPLICANT: Stredler, Wolfgang
TITLE OF INVENTION: PROCESS FOR THE ISOLATION AND
TITLE OF INVENTION: CHARACTERISATION OF A GENE ENZYME SYSTEM FOR INACTIVATION

TITLE OF INVENTION: OF THE HERBICIDE PHENMEDIPHAN AND TRANSFER OF THE GENE
TITLE OF INVENTION: INTO PLANTS TO PRODUCE HERBICIDE-TOLERANT PLANTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
STREET: 1180 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-8403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948.703
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/446.934
FILING DATE:
APPLICATION NUMBER: US 08/196.361
FILING DATE: 02-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/615.448
FILING DATE: 19-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/353.871
FILING DATE: 18-MAY-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P3817384.0
FILING DATE: 19-MAY-1988
ATTORNEY/AGENT INFORMATION:
NAME: Meliman, Edward A.
REGISTRATION NUMBER: 24,735
REFERENCE/DOCKET NUMBER: P/951-114
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-382-0700
TELEFAX: 212-382-0888
TELEX: 236925
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 493 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-948-703-7

Query Match 75.0%; Score 36; DB 3; Length 493;
Best Local Similarity 71.4%; Pred. No. 2.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGCVDG 8
DB 15 GGCLEDG 21

RESULT 6
US-08-426-125-1
Sequence 1, Application US/08426125
Patent No. 5853973
GENERAL INFORMATION:
APPLICANT: Kakefuda, Genichi
APPLICANT: Ott, Karl-Heinz
APPLICANT: Kwagh, Jae-Gyu
APPLICANT: Stockton, Gerald W.
TITLE OF INVENTION: Structure-Based Designed Herbicide
TITLE OF INVENTION: Resistant Products
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby

STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022-7513
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,125
FILING DATE: 20-APR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Robinson, Joseph
REGISTRATION NUMBER: 33,448
REFERENCE/DOCKET NUMBER: 0646/0A674
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)-527-7783
TELEFAX: (212)-753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 599 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: zea mays
US-08-426-125-1

Query Match 70.8% Score 34; DB 2; Length 599;
Best Local Similarity 62.5%; Pred. No. 6.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGCVODG 8
DB 236 GGGCARG 243

RESULT 7
US-08-455-355-1
; Sequence 1, Application US/08455355
; Patent No. 5928937
; GENERAL INFORMATION:
; APPLICANT: Kakefuda, Genichi
; APPLICANT: Olt, Karl-Heinz
; APPLICANT: Kwagh, Jae-Gyu
; APPLICANT: Stockton, Gerald W.
; TITLE OF INVENTION: Structure-Based Designed Herbicide
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022-7513
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,355
FILING DATE: 31-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Robinson, Joseph
REGISTRATION NUMBER: 33,448
REFERENCE/DOCKET NUMBER: 0646/1A674-US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)-527-7783
TELEFAX: (212)-753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 599 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: zea mays
US-08-455-355-1

Query Match 70.8% Score 34; DB 2; Length 599;
Best Local Similarity 62.5%; Pred. No. 6.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGCVODG 8
DB 236 GGGCARG 243

RESULT 8
US-08-321-356-2
; Sequence 2, Application US/08321356
; Patent No. 5633437
; GENERAL INFORMATION:
; APPLICANT: Bernasconi, Paul
; APPLICANT: Woodworth, Allison R.
; TITLE OF INVENTION: GENE EXHIBITING RESISTANCE TO
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SANDOZ AGRO INC.
; STREET: 975 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/321,356
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Marcus-Wyner, Lynn
REGISTRATION NUMBER: 34869
REFERENCE/DOCKET NUMBER: 133-0720
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)354-3588
TELEFAX: (415)857-1125
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 648 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-321-356-2

Query Match 70.8% Score 34; DB 1; Length 648;

Best Local Similarity 62.5%; Pred. No. 6.5e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGGCVQDG 8
||||: 1
Db 285 GGGCLNSG 292

RESULT 9

US-08-321-356-4
; Sequence 4, Application US/08321356
; Patent No. 5633437
; GENERAL INFORMATION:
; APPLICANT: Bernasconi, Paul
; APPLICANT: Woodworth, Allison R.
; TITLE OF INVENTION: GENE EXHIBITING RESISTANCE TO
; TITLE OF INVENTION: ACETOLACTATE SYNTHASE INHIBITOR HERBICIDES
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SANDOX AGRO INC.
; STREET: 975 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321.356
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Marcus-Wyner, Lynn
; REGISTRATION NUMBER: 34869
; REFERENCE/DOCKET NUMBER: 133-0720
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)354-3588
; TELEFAX: (415)857-1125
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 648 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: Protein
; US-08-321-356-4

Query Match 70.8%; Score 34; DB 1; Length 648;
Best Local Similarity 62.5%; Pred. No. 6.5e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGGCVQDG 8
||||: 1
Db 285 GGGCLNSG 292

RESULT 10

US-08-469-412A-7
; Sequence 7, Application US/08469412A
; Patent No. 5856125
; GENERAL INFORMATION:
; APPLICANT: Mavrothalassitis, George J.
; APPLICANT: Blair, Donald G.
; APPLICANT: Fisher, Robert J.
; APPLICANT: Beal Jr., Gregory J.
; APPLICANT: Athanasoulou, Meropi A.
; APPLICANT: Sgouras, Dionysios N.
; TITLE OF INVENTION: The ERF Genetic Locus and Its Products
; NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469.412A

FILING DATE: 05-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Garrett-Mackowski, Eugenia

REGISTRATION NUMBER: 37,330

REFERENCE/DOCKET NUMBER: 015280-229000

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 543 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: Protein

FEATURE:

NAME/KEY: Protein

LOCATION: 1..543

OTHER INFORMATION: /note="murine ERF amino acid sequence

OTHER INFORMATION: (first 8 amino acids from first exon not

OTHER INFORMATION: included)"

US-08-469-412A-7

Query Match 68.8%; Score 33; DB 2; Length 543;
Best Local Similarity 62.5%; Pred. No. 7.7e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGGCVQDG 8
||||: 1
Db 490 GGGCLSG 497

RESULT 11

US-09-021-715-7
; Sequence 7, Application US/09021715
; Patent No. 6194547
; GENERAL INFORMATION:
; APPLICANT: Mavrothalassitis, George J.
; APPLICANT: Blair, Donald G.
; APPLICANT: Fisher, Robert J.
; APPLICANT: Beal Jr., Gregory J.
; APPLICANT: Athanasoulou, Meropi A.
; APPLICANT: Sgouras, Dionysios N.
; TITLE OF INVENTION: The ERF Genetic Locus and Its Products
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/021,715
FILING DATE: 10-Feb-1998
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Garrett-Mackowski, Eugenia
REGISTRATION NUMBER: 37,330
REFERENCE/DOCKET NUMBER: 015280-229000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 543 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..543
OTHER INFORMATION: /note="murine ERF amino acid sequence
(first 8 amino acids from first exon not
included)"
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-021-715-7

Query Match
Best Local Similarity 68.8%; Score 33; DB 4; Length 543;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGGCVDG 8
DB 490 GGGCLSG 497

RESULT 12
US-08-143-365A-3
; Sequence 3, Application US/08143365A
; Patent No. 5726292
; GENERAL INFORMATION:
; APPLICANT: Lowell, George H
; TITLE OF INVENTION: IMMUNO-POTENTIATING SYSTEMS FOR
; NUMBER OF SEQUENCES: 15
; PREPARATION OF IMMUNOGENIC MATERIALS
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hendricks and Assoc.
; STREET: 9669 A Main Street
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/143,365A
; FILING DATE: 29-Oct-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna M
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: Lowell-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 425-4250
; TELEFAX: (703) 425-2767
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
```

```

STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
US-08-143-365A-3

Query Match
Best Local Similarity 66.7%; Score 32; DB 1; Length 14;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCVDG 6
DB 2 GGGCTG 7

RESULT 13
US-08-143-365A-10
; Sequence 10, Application US/08143365A
; Patent No. 5726292
; GENERAL INFORMATION:
; APPLICANT: Lowell, George H
; TITLE OF INVENTION: IMMUNO-POTENTIATING SYSTEMS FOR
; NUMBER OF SEQUENCES: 15
; PREPARATION OF IMMUNOGENIC MATERIALS
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hendricks and Assoc.
; STREET: 9669 A Main Street
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/143,365A
; FILING DATE: 29-Oct-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna M
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: Lowell-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 425-4250
; TELEFAX: (703) 425-2767
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; MOLECULE TYPE: unknown
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
US-08-143-365A-10

Query Match
Best Local Similarity 66.7%; Score 32; DB 1; Length 19;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCVDG 6
DB 7 GGGCTG 12

RESULT 14
```

US-08-318-157B-2
: Sequence 2, Application US/08318157B
: Patent No. 5874540
: GENERAL INFORMATION:
: APPLICANT: HANSEN, Hans J.
: APPLICANT: ARMOUR, Kathryn L.
: TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
: TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES
: NUMBER OF SEQUENCES: 58
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W., Suite 500
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20007-5109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/318,157B
: FILING DATE: 05-OCT-1994
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: SAYE, Bernhard D.
: REGISTRATION NUMBER: 28,665
: REFERENCE/DOCKET NUMBER: 18733/464
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)672-5300
: TELEFAX: (202)672-5399
: TELEX: 904136
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 119 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-318-157B-2

Query Match 66.7%; Score 32; DB 2; Length 119;
Best Local Similarity 75.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCGCVQDG 8
||| |||
DB 8 GCGLVQSG 15

RESULT 15
US-08-471-570-2
: Sequence 2, Application US/08471570
: Patent No. 5750371
: GENERAL INFORMATION:
: APPLICANT: IGARASHI, Koichi
: APPLICANT: SENOO, Masaharu
: APPLICANT: WATANABE, Tatsuya
: TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
: NUMBER OF SEQUENCES: 18
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &
: STREET: 130 Water Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: US
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/471,570
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/149,664
: FILING DATE:
: APPLICATION NUMBER: US 07/743369
: FILING DATE: 16-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: LINEK, Ernest V.
: REGISTRATION NUMBER: 29822
: REFERENCE/DOCKET NUMBER: 40897
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)523-3400
: TELEFAX: (617)523-6440
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 201 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-471-570-2

Query Match 66.7%; Score 32; DB 1; Length 201;
Best Local Similarity 71.4%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GCGCVQDG 8
||| |||
DB 111 GACTQDGG 117

Search completed: June 13, 2001, 14:27:06
Job time: 627 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:25:44 ; Search time 150.28 Seconds
(without alignments)
3.043 Million cell updates/sec

Title: PCT-US01-05825A-23

Perfect score: 44
Sequence: 1 GGGVQPG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues
Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /_Geneseq_0401:*
2: /SIDS6/gcgdata/geneseq/geneseq/AA1980.DAT:*
3: /SIDS6/gcgdata/geneseq/geneseq/AA1981.DAT:*
4: /SIDS6/gcgdata/geneseq/geneseq/AA1982.DAT:*
5: /SIDS6/gcgdata/geneseq/geneseq/AA1983.DAT:*
6: /SIDS6/gcgdata/geneseq/geneseq/AA1984.DAT:*
7: /SIDS6/gcgdata/geneseq/geneseq/AA1985.DAT:*
8: /SIDS6/gcgdata/geneseq/geneseq/AA1986.DAT:*
9: /SIDS6/gcgdata/geneseq/geneseq/AA1987.DAT:*
10: /SIDS6/gcgdata/geneseq/geneseq/AA1988.DAT:*
11: /SIDS6/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SIDS6/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SIDS6/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SIDS6/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SIDS6/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SIDS6/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SIDS6/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SIDS6/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SIDS6/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SIDS6/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SIDS6/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SIDS6/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	44	100.0	8 21	Y79127 Peptide antagonist
2	44	100.0	13 20	W86096 Peptide from human
3	44	100.0	13 20	W86076 Peptide from human
4	44	100.0	13 20	W86088 Peptide from human
5	44	100.0	18 16	W73953 IgG heavy chain fr
6	44	100.0	18 16	R82833 N-terminal of C179
7	44	100.0	20 20	W94487 Human adult heart
8	44	100.0	20 21	Y79130 Human adult heart
9	44	100.0	30 17	R87049 Human group III he
10	44	100.0	30 21	Y68811 Human heavy chain
11	44	100.0	40 7	P61028 H-chain variable r

12	44	100.0	62 16	R76969 HSV-neutralising a
13	44	100.0	73 21	Y64774 Human 5' EST relat
14	44	100.0	84 21	Y64925 Human 5' EST relat
15	44	100.0	87 17	R92997 Homologous sequenc
16	44	100.0	93 20	Y11665 Human 5' EST seque
17	44	100.0	96 21	Y56646 Partial peptide fr
18	44	100.0	96 21	Y64844 Human 5' EST relat
19	44	100.0	97 21	B40089 Anti-hiL12 antiod
20	44	100.0	97 21	B40092 Anti-hiL12 antiod
21	44	100.0	97 21	B40112 Anti-hiL12 antiod
22	44	100.0	97 21	B40135 Anti-hiL12 antiod
23	44	100.0	98 14	R34279 Human TNF binding
24	44	100.0	98 15	R52066 Heavy chain variab
25	44	100.0	98 15	R54816 SPA-reactive VH re
26	44	100.0	98 16	R72074 DP54 VH region. H
27	44	100.0	98 19	W59614 Anti-RSV F protein
28	44	100.0	98 21	B40084 Anti-hiL12 antiod
29	44	100.0	98 21	B40087 Anti-hiL12 antiod
30	44	100.0	98 21	B40085 Anti-hiL12 antiod
31	44	100.0	98 21	B40086 Anti-hiL12 antiod
32	44	100.0	98 21	B40087 Anti-hiL12 antiod
33	44	100.0	98 21	B40093 Anti-hiL12 antiod
34	44	100.0	98 21	B40094 Anti-hiL12 antiod
35	44	100.0	98 21	B40095 Anti-hiL12 antiod
36	44	100.0	98 21	B40096 Anti-hiL12 antiod
37	44	100.0	98 21	B40097 Anti-hiL12 antiod
38	44	100.0	98 21	B40107 Anti-hiL12 antiod
39	44	100.0	98 21	B40108 Anti-hiL12 antiod
40	44	100.0	98 21	B40113 Anti-hiL12 antiod
41	44	100.0	98 21	B40130 Anti-hiL12 antiod
42	44	100.0	98 21	B40134 Anti-hiL12 antiod
43	44	100.0	98 21	B40136 Anti-hiL12 antiod
44	44	100.0	98 21	B40137 Partial peptide fr
45	44	100.0	98 21	Y56660

ALIGNMENTS

RESULT 1
ID Y79127 standard; Peptide: 8 AA.
XX Y79127;
AC 05-JUN-2000 (first entry)
DT
XX
XX
DE Peptide antagonist of zonulin.
KW Zonulin; antagonist; zonula occludens toxin receptor;
KW blood-brain barrier; antiinflammatory; cerebroprotective;
KW neuroprotective; dermatological; antiulcer; antiviral;
KW antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
KW gastrointestinal inflammation; therapy.
XX
XX Synthetic.
OS
XX
XX WO200007609-A1.
PN
XX
XX 17-FEB-2000.
PD
XX
XX 28-JUL-1999; 99WO-US16683.
PF
XX
XX 03-AUG-1998; 98US-0127815.
PR
XX
XX (UYWA-) UNIV MARYLAND BALTIMORE.
PA
XX
XX Fasano A;
PI
XX
XX WPI, 2000-205565/18.
DR
XX
XX New peptide antagonist of zonulin useful as antiinflammatory agent for

	Query March	100.0%	Score 44;	DB 21;	length 8;
Best Local Similarity		100.0%	Pred. No. 3.2e+05;		
Matches	8; Conservative	0;	Mismatches	0;	Indels
Oy	1 GGGIvOPG 8				0;
Db	1 999IvOPG 8				

```

PI Carr FJ;
XX
DR WPI; 1999-04S301/04.
XX
PR Reducing immunogenicity of proteins - by modifying the amino acid
PT sequence of the protein to eliminate potential epitopes for T-cells
PS of a given species
XX
PS Example 5; Page 34; 77pp; English.
XX
CC The invention relates to a method for the production of non-immunogenic
CC proteins. The method comprises determining at least part of the amino
CC acid sequence of the protein; (b) identifying in the amino acid sequence
CC one or more potential epitopes for T-cells (T-cell epitopes) of the given
CC species; and (c) modifying the amino acid sequence to eliminate at least
CC one of the T-cell epitopes identified in step (b) thereby to eliminate or
CC reduce the immunogenicity of the protein when exposed to the immune
CC system of the given species. A method of analysing a pre-existing protein
CC to predict the basis for immunogenic responses is also provided. The
CC methods can be used particularly for reducing the immunogenicity of
CC immunoglobulins or therapeutic proteins, e.g. Streptokinase. The products
CC can be used for diagnosis and therapy. Sequences W86075-W86114 represent
CC peptides from humanised A33 heavy chain variable region predicted by
CC peptide threading.
XX
SQ Sequence 13 AA;

Query Match          100.0%; Score 44; DB 20; Length 13;
Best Local Similarity 100.0%; Prd. No. 0.2;
Matches      8; Conservative    0; Mismatches     0; Indels    0; Gaps    0;

QY      1 GGGLVQPQG 8
        |||||
Db       6 ggglvqpg 13

RESULT   3
W86076
ID W86076 standard; peptide; 13 AA.
XX
AC W86076;
XX
DT 03-MAR-1999 (first entry)
XX
DE Peptide from humanised A33 heavy chain variable region.
XX
KW Non-immunogenic; epitope; T-cell; immunogenicity; immune system;
KW immunoglobulin; therapeutic; streptokinase; diagnosis; human; A33.
OS Homo sapiens.
XX
FN WO9852976-A1.
XX
PD 26-NOV-1998.
XX
PF 21-MAY-1998; 98WO-GB01473.
XX
PR 14-APR-1998; 98GB-0007751.
PR 21-MAY-1997; 97GB-0010480.
PR 31-JUL-1997; 97GB-0016197.
PR 28-NOV-1997; 97GB-0025270.
PR 02-DEC-1997; 97US-0067235.
XX
PA (BIOV-) BIOVATION LTD.
XX
PI Carr FJ;
XX
DR WPI; 1999-04S301/04.
XX
PR Reducing immunogenicity of proteins - by modifying the amino acid
PT sequence of the protein to eliminate potential epitopes for T-cells
of a given species

```

XX Example 5; Page 33; 77pp; English.

PS The invention relates to a method for the production of non-immunogenic

XX proteins. The method comprises determining at least part of the amino

CC acid sequence of the protein; (b) identifying in the amino acid sequence

CC one or more potential epitopes for T-cells (T-cell epitopes) of the given

CC species; and (c) modifying the amino acid sequence to eliminate at least

CC one of the T-cell epitopes identified in step (b) thereby to eliminate or

CC reduce the immunogenicity of the protein when exposed to the immune

CC system of the given species. A method of analysing a pre-existing protein

CC to predict the basis for immunogenic responses is also provided. The

CC methods can be used particularly for reducing the immunogenicity of

CC immunoglobulins or therapeutic proteins, e.g. Streptokinase. The products

CC can be used for diagnosis and therapy. Sequences W86075-W86114 represent

CC peptides from humanised A33 heavy chain variable region predicted by

CC peptide threading.

SO Sequence 13 AA:

QY 1 GGGVQPG 8

DB 6 gggivqpg 13

RESULT 4

W86088

ID W86088 standard; peptide; 13 AA.

XX W86088;

AC 03-MAR-1999 (first entry)

DT

XX Peptide from humanised A33 heavy chain variable region.

DE

XX Non-immunogenic; epitope; T-cell; immunogenicity; immune system;

KW Immunoglobulin; therapeutic; streptokinase; diagnosis; human; A33.

KM

XX Homo sapiens.

OS

XX W09852976-A1.

PN

XX 26-NOV-1998.

PD

XX 21-MAY-1998; 98WO-GB01473.

PE

XX 14-APR-1998; 98GB-0007751.

PR

XX 21-MAY-1997; 97GB-0010480.

PR

XX 31-JUL-1997; 97GB-0016197.

PR

XX 28-NOV-1997; 97GB-0025270.

PR

XX 02-DEC-1997; 97US-0067235.

PA

XX (BIOV-) BIOVATION LTD.

PI

XX Carr FJ;

PT WPI; 1999-045301/04.

DR

XX Reducing immunogenicity of proteins - by modifying the amino acid

PT sequence of the protein to eliminate potential epitopes for T-cells

XX of a given species

XX Example 5; Page 33; 77pp; English.

PS The invention relates to a method for the production of non-immunogenic

CC proteins. The method comprises determining at least part of the amino

CC acid sequence of the protein; (b) identifying in the amino acid sequence

CC one or more potential epitopes for T-cells (T-cell epitopes) of the given

CC species; and (c) modifying the amino acid sequence to eliminate at least

CC one of the T-cell epitopes identified in step (b) thereby to eliminate or

CC reduce the immunogenicity of the protein when exposed to the immune

CC system of the given species. A method of analysing a pre-existing protein

CC to predict the basis for immunogenic responses is also provided. The

CC methods can be used particularly for reducing the immunogenicity of

CC immunoglobulins or therapeutic proteins, e.g. Streptokinase. The products

CC can be used for diagnosis and therapy. Sequences W86075-W86114 represent

CC peptides from humanised A33 heavy chain variable region predicted by

CC peptide threading.

SO Sequence 13 AA:

QY 1 GGGVQPG 8

DB 3 gggivqpg 10

RESULT 5

W73953

ID W73953 standard; peptide; 18 AA.

XX W73953;

AC 26-APR-1999 (first entry)

DT

XX IgG heavy chain fragment, IgG (VH) (18).

DE

XX IgG heavy chain; immune complex removal; conglutinin; infection;

KW serum sickness; chronic immune complex nephritis; parasitic disease;

KM autoimmune disease; neoplastic disease; immune pathogenesis; therapy.

KX

XX Synthetic.

OS

XX W09506254-A1.

PN

XX 02-MAR-1995.

PD

XX 24-AUG-1994; 94WO-US09407.

PE

XX 24-AUG-1993; 93US-0111071.

PR

XX (IMMU-) APPLIED IMMUNE SCI INC.

PA

XX Lee YM, Odonoghue G, Okarma TB;

PI

XX WPI; 1995-106945/14.

DR

XX New biomedical device for binding immune complexes - comprising

PT conglutinin covalently bound to a solid phase material

PT

XX Example 4; Fig 11; 58pp; English.

PS This sequence represents a fragment of the IgG heavy

CC chain. This sequence was isolated using the biomedical device of the

CC invention, which comprises: (a) a biocompatible support; and

CC (b) conglutinin which is covalently bound to the support. The device can

CC be used for removing immune complexes (IC) from a fluid. This can be used

CC for treating a mammal or for evaluating the status of a mammal. It can be

CC used for treating diseases such as serum sickness, chronic immune complex

CC nephritis, bacterial infections, viral infections, parasitic diseases,

CC autoimmune diseases or neoplastic diseases. The devices can also be used

CC for purification of IC and to study immune pathogenesis and antigens

CC associated with a disease state. The covalently immobilised conglutinin

CC can capture IC with high capacity and specificity and allows elution

CC under mild conditions. The devices are stable with no loss of IC binding

CC activity after repeated use.

SQ Sequence 18 AA;

Query Match 100.0%; Score 44; DB 16; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGIVQPG 8
| | | | | | | |
Db 8 99gIvqpg 15

RESULT 6
R82833
ID R82833 standard; Peptide: 18 AA.

AC R82833;
DT 09-MAY-1996 (first entry)
DE N-terminal of C179 heavy chain.

KM Antibody; human; Influenza type A virus; H1N1; H2N2; H3N2; HA; diagnosis;
KW haemagglutinin; variable light chain; therapy; CDR;
XX complementarity determining region.

OS Mus musculus.

PN EP675199-A2.

PD 04-OCT-1995.

PF 14-MAR-1995; 95EP-0301664.

PR 30-MAR-1994; 94JP-0082693.

PA (TAKI) TAKARA SHUZO CO LTD.

PI Atsushi O, Hirofumi Y, Ikunoshin K, Takashi T, Yoshinobu O;

DR WPI; 1995-338286/44.

PT Gene encoding variable region of anti-human influenza A type virus
antibody - useful for prodn. of artificial antibodies

PS Example 2; Page 24; 42pp; English.

CC R82833 and R82834 represent the N-terminal sequences of the heavy chain
CC and light chain of an anti-human influenza A type virus antibody. The
CC antibody these sequences were obtained from was isolated from hybridoma
CC C179. The antibody recognises the stem region of the haemagglutinin (HA)
CC molecule of the H1N1 and H2N2 subtypes of human influenza A type virus.
CC and shows neutralisation activity against these two subtypes. The
CC antibody shows no recognition of the H3N2 subtype. The DNA encoding the
CC variable heavy chain sequence (see T04159) and the variable light chain
CC coding sequence (see T04160) are useful in the production of artificial
CC antibodies and antigen-binding polypeptides. These artificial antibodies
CC and polypeptides are useful in the diagnosis and treatment of human
CC influenza. As the antibodies recognise the stem region of the HA
CC molecule, the influenza virus will be recognised even if the HA molecule
CC changes. This provides an advantage over current vaccines, as the virus
CC periodically alters its HA molecule.

SO Sequence 18 AA;

Query Match 100.0%; Score 44; DB 16; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGIVQPG 8
| | | | | | | |
Db 8 99gIvqpg 15

RESULT 7

W94487
ID W94487 standard; peptide: 20 AA.

AC W94487;

DT 21-APR-1999 (first entry)

DE Human adult heart zonulin N-terminal peptide.

KW Zonulin; mammalian tight junction; zonula occludens toxin; ZOT;
KW Vibrio cholerae; vaccine; cholera toxin; polyclonal antibody;
KW intestinal mucosa; nasal mucosa; blood brain barrier.

OS Homo sapiens.

PN W09852415-A1.

PD 26-NOV-1998.

PF 28-APR-1998; 98WO-US07636.

PR 21-MAY-1997; 97US-0859931.

PA (UYMA-) UNIV MARYLAND BALTIMORE.

PI Fasano A;

DR WPI; 1999-070123/06.

PT New purified zonulin - which is capable of reversibly opening
mammalian tight junctions, used for enhancing the delivery of agents
PT across intestinal and nasal mucosa and blood brain barrier
PS Claim 2; Page 45; 64pp; English.

CC The present invention describes pure zonulin which has an apparent
CC molecular weight of 47 kD, as determined by SDS-PAGE, which is
CC recognised by both anti-tau polyclonal antibody and by anti-zonula
CC occludens toxin (ZOT) polyclonal antibody, and is capable of reversibly
CC opening mammalian tight junctions. Zonulin proteins function as
CC physiological modulators of mammalian tight junctions. They can be used
CC for enhancing the absorption of therapeutic agents across tight
CC junctions of intestinal and nasal mucosa and across tight junctions of
CC the blood brain barrier. Zonulin can be used with agents such as drugs,
CC e.g. lidocaine, adenosine, dobutamine, dopamine, epinephrine,
CC norepinephrine, phenolamine, doxapram, alfentanil, dezocin, nalbuphine,
CC buprenorphine, naloxone, ketorolac, midazolam, propofol, metacurline,
CC mivacurium, succinylcholine, cytarabine, mitomycin doxorubicin,
CC vincristine, vinblastine, methicillin, mezlocillin, piperacillin,
CC cefoxitin, ceftriaxone, cefmetazole and aztreonam, a hormone e.g.
CC testosterone, nandrolone, menotropins, insulin, urofollitropin,
CC interferon-alpha, interferon-beta, interferon-gamma, interleukin-1
CC (IL-1), IL-2, IL-4, IL-8, polyvalent IgG, specific IgG, IgA, or IgM.
CC The proteins can also be used for the production of antibodies which can
CC be used to assay for zonulin in body tissue or fluids, or in affinity-
CC purification of zonulin. The present sequence represents an N-terminal
CC peptide of zonulin.

SO Sequence 20 AA;

Query Match 100.0%; Score 44; DB 20; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGIVQPG 8
| | | | | | | |
Db 8 99gIvqpg 15


```
RESULT 8
Y79130
ID Y79130 standard; Peptide: 20 AA.
XX
AC Y79130;
XX
DE 05-JUN-2000 (first entry)
XX
DE Human adult heart zonulin N-terminal sequence.
XX
KW Zonulin; antagonist; zonula occludens toxin receptor;
KW human; blood-brain barrier; antiinflammatory;
KW gastrointestinal inflammation; therapy.
XX
OS Homo sapiens.
XX
PN WO200007609-A1.
XX
PD 17-FEB-2000.
XX
PE 28-JUL-1999; 99WO-US16683.
XX
PR 03-AUG-1998; 98US-0127815.
XX
PA (UYMA-) UNIV MARYLAND BALTIMORE.
XX
PI Fasano A;
XX
DR WPI: 2000-205565/18.
XX
PT New peptide antagonist of zonulin useful as antiinflammatory agent for
PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
XX
PS Example 3; Fig 6; 69pp; English.
XX
CC The present sequence is that of the N-terminal region of adult
CC human heart zonulin. The N-terminal sequences of human adult and
CC foetal zonulins (see Y79130-36) were compared with Vibrio cholerae
CC zonula occludens toxin (ZOT) to identify a common motif thought
CC to be involved in receptor binding. Peptide antagonists (see
CC Y79105-29) based on this motif are useful as antiinflammatory
CC agents for treatment of gastrointestinal inflammation, and for
CC treatment of conditions associated with breakdown of the blood-brain
CC barrier.
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 44; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGIVQPG 8
   |||||||
DB 8 9991VQPG 15

RESULT 9
R87049
ID R87049 standard; Peptide: 30 AA.
XX
AC R87049;
XX
DE 25-JUN-1996 (first entry)
XX
DE Human group III heavy chain framework 1.
XX
KW Humanised antibody; interleukin-5; IL-5; recombinant antibody;
KW antibody engineering; monoclonal antibody; Mab; 39D10; CDR;
KW complementarity determining region; heavy chain; framework;
KW eosinophilla; allergy; asthma.
XX
```

```
OS Homo sapiens.
XX
PN WO9535375-A1.
XX
PD 28-DEC-1995.
XX
PE 16-JUN-1995; 95WO-GB01411.
XX
PR 17-JUN-1994; 94GB-0012230.
XX
PA (CLLT ) CELLTech THERAPEUTICS LTD.
XX
PI Athwal DS, Bodmer MW, Emtage JS;
XX
DR WPI: 1996-058412/06.
XX
PT Anti-human IL-5 recombinant antibody - useful for preventing or
PT reducing eosinophilla and for treating certain allergic diseases,
PT esp. asthma
XX
PS Example 3; Fig 4; 69pp; English.
XX
CC Framework regions (R87049-52) of human group III (gp3) germ line
CC antibody heavy chain showed homology to corresponding regions
CC (R87053-56, respectively) of the rat anti-human interleukin-5
CC monoclonal antibody 39D10 heavy chain (see R87039). This homology
CC was utilised in the prodn. of a humanised 39D10 VH (R87058) in
CC which rat 39D10 VH complementarity determining regions were grafted
CC into the human gp3 framework.
XX
SQ Sequence 30 AA;

Query Match 100.0%; Score 44; DB 17; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGIVQPG 8
   |||||||
DB 8 9991VQPG 15

RESULT 10
Y68811
ID Y68811 standard; Peptide: 30 AA.
XX
AC Y68811;
XX
DE 16-MAY-2000 (first entry)
XX
DE Human heavy chain variable domain framework sequence without CDRs.
XX
KW Complementarity determining region; CDR; rat; immunoglobulin G; IgG;
KW CD3 antigen complex; chimeric antibody; immunosuppression; heavy chain;
KW CD3 antigen-T-cell receptor complex; graft rejection; cancer;
KW framework region.
XX
OS Homo sapiens.
XX
PN WO200005268-A1.
XX
PD 03-FEB-2000.
XX
PE 21-JUL-1999; 99WO-GB02380.
XX
PR 21-JUL-1998; 98GB-0015909.
XX
PA (BIOI-) BIO INT LTD.
XX
PI Waldmann H, Frewin M;
XX
DR WPI: 2000-182655/16.
XX
```

PT New humanised anti-CD3 antibodies, used for treating cancer or for
 PT immunosuppression and preventing graft rejection
 PS Claim 9: Page 47; 56pp; English.
 XX
 CC Y68811-14 represent human heavy chain variable domain framework
 CC sequences without complementarity determining regions (CDRs). The
 CC sequences are used to produce the antibodies of the invention. The
 CC specification describes chimeric human/rodent anti-CD3 antibodies,
 CC which have a rodent anti-CD3 immunoglobulin G (19g) light chain
 CC variable region and a human heavy chain variable region. The
 CC anti-CD3 antibodies can render T-cells non-functional by antibody
 CC blockade of the CD3 antigen-T-cell receptor (TCR) complex. They can be
 CC used for immunosuppression, particularly for the control of graft
 CC rejection. The antibodies can also enhance or re-direct T-cell
 CC responses to antigens. They can be used in the treatment of cancer.
 XX
 SO Sequence 30 AA;

Query Match 100.0%; Score 44; DB 21; Length 30;
 Best Local Similarity 100.0%; Pred. No. 0.46; 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0;

OY 1 GGGLVQPG 8
 |||||
 DB 8 gggglvqpg 15

RESULT 11
 P61028
 ID P61028 standard; Protein; 40 AA.

AC P61028;
 XX
 DT 21-OCR-1991 (first entry)
 XX
 DE H-chain variable region of plasminogen activator antibody.
 XX
 KW Mab; anti-plasminogen activator; PA.

XX JF61172900-A.
 XX
 PD 04-AUG-1986.
 XX
 PF 29-JAN-1985; 85JP-0013399.
 XX
 PR 29-JAN-1985; 85JP-0013399.

XX (ASAH) ASAH CHEMICAL IND KK.

XX WPI; 1986-242924/37.

PT Monoclonal: antibody obtd. from hybrid cells - used in
 PT Immune assay of plasminogen activator tissue obtd. e.g. from
 PT human embryo pulmonary cells w.r.t.

PS Claim 1; Page 849; 9pp; Japanese.

XX Monoclonal antibodies raised to the plasminogen activator antigen
 CC may be used in isolation or detection of pure PA.
 CC
 XX Sequence 40 AA;

Query Match 100.0%; Score 44; DB 7; Length 40;
 Best Local Similarity 100.0%; Pred. No. 0.61; 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0;

OY 1 GGGLVQPG 8
 |||||
 DB 8 gggglvqpg 15

RESULT 12
 R76969
 ID R76969 standard; peptide; 62 AA.

AC R76969;
 XX
 DT 23-FEB-1996 (first entry)

DE HSV-neutralising antibody clone Fab fragment heavy chain variable domain.

XX Herpes simplex virus; type I; type II; monoclonal antibody;
 KW diagnosis; neutralisation; immunotherapy.
 XX
 OS Homo sapiens.

XX W09518634-A1.
 PN
 PD 13-JUL-1995.

XX 04-JAN-1995; 95MO-US00067.
 PF
 PR 04-JAN-1994; 94US-0178201.

XX (SCRI) SCRIPPS RES INST.
 XX
 PI Burtonl R, Burton DR, Sanna PP, Williamson RA;

XX WPI; 1995-254909/33.

PT Human monoclonal antibodies that neutralise Herpes simplex virus
 PT (HSV) types 1 and 2 - used for diagnosis and passive immuno-therapy
 PT of HSV infections.
 XX
 PS Example 4; Page 60; 100pp; English.

XX R76965-R76579 are the Fab fragment heavy chain variable domains of
 CC human anti-herpes monoclonal antibodies. These antibodies are capable
 CC of neutralising both herpes simplex virus (HSV) types I and II by
 CC binding an epitope present on glycoprotein D. The antibody may be used
 CC for detecting HSV in vivo or in vitro; for passive immunotherapy (pret.
 CC prophylactically) of HSV infection (eg. genital, oral or ocular herpes),
 CC partic. as its Fab fragment and as a competitive reagent for detecting
 CC neutralising anti-HSV antibodies in a sample. Anti-idiotypic antibodies
 CC raised against the mab can be used for active immunotherapy of HSV
 CC infection.
 CC
 XX Sequence 62 AA;

Query Match 100.0%; Score 44; DB 16; Length 62;
 Best Local Similarity 100.0%; Pred. No. 0.93; 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0;

OY 1 GGGLVQPG 8
 |||||
 DB 4 gggglvqpg 11

RESULT 13
 Y64774
 ID Y64774 standard; Protein; 73 AA.

AC Y64774;
 XX
 DT 01-FEB-2000 (first entry)

DE Human 5' EST related polypeptide SEQ ID NO:935.

XX Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
 KW gene therapy; chromosome mapping; upstream regulatory sequence;
 KW forensic; location; development; protein synthesis; stability;
 KW regulation; identification.

```

XX OS Homo sapiens.
XX PN WO953051-A2.
XX PD 21-OCT-1999.
XX PE 09-APR-1999; 99MO-IB00712.
XX PR 09-APR-1998; 98US-0057719.
XX PR 28-APR-1998; 98US-0069047.
XX PA (GEST ) GENSET.
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX DR WPI; 2000-038446/03.
XX DR N-PSDB; 242388.
XX PT Novel secreted protein 5' expressed sequence tag sequences used in
XX PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
XX PS Claim 3; Page 638; 837pp; English.
XX CC 242265 to 243075 represent novel 5' expressed sequence tag (EST)
XX CC sequences, corresponding to human secreted proteins. Y64651 to Y65438
XX CC represent the EST-related proteins corresponding to 242265 to 243052.
XX CC The 5' ESTs can be used for producing secreted human gene products.
XX CC They can be used to identify and isolate 5' untranslated regions (UTRs)
XX CC and upstream regulatory regions which control the location, development
XX CC stage, rate, and quantity of protein synthesis, as well as stability of
XX CC mRNA. The ESTs are also useful as probes for chromosome mapping, and to
XX CC obtain full length cDNA clones. The ESTs can also be used in forensic
XX CC procedures to identify individuals, or in diagnostic procedures to
XX CC identify individuals having genetic diseases resulting from abnormal
XX CC gene expression. The products may also be used in gene therapy protocols.
XX CC The nucleic acids encoding signal peptides can be used for directing
XX CC extracellular secretion of a polypeptide or the insertion of a
XX CC polypeptide into a membrane, or importing a polypeptide into a cell.
XX CC The proteins encoded by the EST sequences may be useful in treating a
XX CC variety of human conditions. Secreted proteins have therapeutic value,
XX CC and the identification of new secreted proteins is valuable. 242249 to
XX CC 242264 and Y64644 to Y64650 represent sequences used in the
XX CC exemplification of the present invention.
XX SQ Sequence 73 AA:

Query Match          100.0%; Score 44; DB 21; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGLVQPG 8
   |||||
Db 27 ggglvpg 34

RESULT 14
ID Y64925
XX Y64925 standard; Protein; 84 AA.
XX AC Y64925;
XX DT 01-FEB-2000 (first entry)
XX DE Human 5' EST related polypeptide SEQ ID NO:1086.
XX KM Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
XX KM gene therapy; chromosome mapping; upstream regulatory sequence;
XX KM forensic; location; development; protein synthesis; stability;
XX KM regulation; identification.
XX OS Homo sapiens.

```

```

XX PN WO953051-A2.
XX PD 21-OCT-1999.
XX PE 09-APR-1999; 99MO-IB00712.
XX PR 09-APR-1998; 98US-0057719.
XX PR 28-APR-1998; 98US-0069047.
XX PA (GEST ) GENSET.
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX DR WPI; 2000-038446/03.
XX DR N-PSDB; 242539.
XX PT Novel secreted protein 5' expressed sequence tag sequences used in
XX PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
XX PS Claim 3; Page 662-683; 837pp; English.
XX CC 242265 to 243075 represent novel 5' expressed sequence tag (EST)
XX CC sequences, corresponding to human secreted proteins. Y64651 to Y65438
XX CC represent the EST-related proteins corresponding to 242265 to 243052.
XX CC The 5' ESTs can be used for producing secreted human gene products.
XX CC They can be used to identify and isolate 5' untranslated regions (UTRs)
XX CC and upstream regulatory regions which control the location, development
XX CC stage, rate, and quantity of protein synthesis, as well as stability of
XX CC mRNA. The ESTs are also useful as probes for chromosome mapping, and to
XX CC obtain full length cDNA clones. The ESTs can also be used in forensic
XX CC procedures to identify individuals, or in diagnostic procedures to
XX CC identify individuals having genetic diseases resulting from abnormal
XX CC gene expression. The products may also be used in gene therapy protocols.
XX CC The nucleic acids encoding signal peptides can be used for directing
XX CC extracellular secretion of a polypeptide or the insertion of a
XX CC polypeptide into a membrane, or importing a polypeptide into a cell.
XX CC The proteins encoded by the EST sequences may be useful in treating a
XX CC variety of human conditions. Secreted proteins have therapeutic value,
XX CC and the identification of new secreted proteins is valuable. 242249 to
XX CC 242264 and Y64644 to Y64650 represent sequences used in the
XX CC exemplification of the present invention.
XX SQ Sequence 84 AA:

Query Match          100.0%; Score 44; DB 21; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGLVQPG 8
   |||||
Db 27 ggglvpg 34

RESULT 15
ID R92997
XX R92997 standard; Protein; 87 AA.
XX AC R92997;
XX DT 18-MAY-1996 (first entry)
XX DE Homologous sequences to antibody IOR-T1 variable region heavy chain.
XX KM IOR-T1; monoclonal antibody; human; mouse; heavy chain; homology;
XX KM variable region; framework; cloning; computer; algorithm;
XX KM immunogenicity; site-directed mutagenesis; T-lymphocyte epitope;
XX KM tertiary structure; point mutation; antibody engineering;
XX KM protein engineering; humanised antibody; antitumour; cancer; therapy.
XX OS Homo sapiens.

```

```

FH      Key      Location/Qualifiers
FT      Region   26..30
FT      Region   /note= "Amino acids involved in tertiary structure"
FT      Region   50
FT      Region   /note= "Amino acid involved in tertiary structure"
FT      Region   76
FT      Region   /note= "Amino acid involved in tertiary structure"
XX
XX      EP699755-A2.
XX
XX      06-MAR-1996.
XX
XX      27-JUN-1995; 95EP-0201752.
XX
XX      30-JUN-1994; 94CU-0000080.
XX
XX      (IMMG-) CENT IMMUNOLOGIA MOLECULAR.
XX
XX      Mateo de Acosta del Rio CM, Rodriguez RP, Valladares JL;
XX
XX      MPI; 1996-130770/14.
XX
XX      Identifying interspecies differences in amino acid sequence of Ig
XX      T-cell epitopes - by sequence comparison, also humanised antibodies
XX      contg. altered T-cell epitopes, retaining antigen specificity but
XX      not immunogenicity, esp. for tumour treatment
XX
XX      Claim 18; Fig 7; 33pp; English.
XX
XX      The sequence represents residues from a human immunoglobulin with
XX      homology to the heavy chain variable region from mouse monoclonal
XX      antibody 10R-T1 (R92996). The sequence is partial, and
XX      complementarily determining regions are omitted. The sequence is
XX      isolated by comparison of human and mouse immunoglobulins and
XX      analysis for T-lymphocyte antigenic sequences using a computer
XX      algorithm. Residues not within a complementarily determining region,
XX      canonical structure or Verner zone may be modified to reduce
XX      immunogenicity in humans (e.g. in sequence R92996). This method,
XX      which involves the introduction of only a few point mutations into
XX      T-cell epitope coding regions, is generally applicable in humanisation
XX      of mouse antibodies. The resulting humanised antibodies may be used
XX      e.g. as antitumour agents. They retain the antigen recognition of the
XX      original antibody, but are not immunogenic in humans.
XX
XX      Sequence 87 AA:
SQ

```

```

Query Match      100.0%; Score 44; DB 17; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GGGVQPG 8
        |||||
Db      8 gggvqpg 15

```

```

Search completed: June 13, 2001, 14:25:44
Job time: 663 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:23:10 ; Search time 87.97 Seconds
(without alignments)
6.250 Million cell updates/sec

Title: PCT-US01-05825A-23

Perfect score: 44

Sequence: 1 GGGLVQPG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR.67:*

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	36	2 D31485	Ig heavy chain V r
2	44	100.0	38	2 S33402	Ig heavy chain V r
3	44	100.0	59	2 A27606	Ig heavy chain V-a
4	44	100.0	59	2 S36381	Ig heavy chain V r
5	44	100.0	60	2 S36382	Ig heavy chain V r
6	44	100.0	82	2 C34964	Ig heavy chain pre
7	44	100.0	94	2 D25913	Ig heavy chain V r
8	44	100.0	94	2 I67528	CD33 antigen homol
9	44	100.0	95	2 I67527	CD33 antigen homol
10	44	100.0	97	2 PH0872	Ig heavy chain V r
11	44	100.0	97	2 S26885	Ig heavy chain V r
12	44	100.0	97	2 S26886	Ig heavy chain V r
13	44	100.0	97	2 S26890	Ig heavy chain V r
14	44	100.0	97	2 S46462	Ig heavy chain V r
15	44	100.0	97	2 S26895	Ig heavy chain V r
16	44	100.0	97	2 S54855	Ig heavy chain V r
17	44	100.0	98	2 PI0121	Ig heavy chain V-I
18	44	100.0	98	2 PH0874	Ig heavy chain V r
19	44	100.0	98	2 PI0123	Ig heavy chain V-I
20	44	100.0	98	2 S26896	Ig heavy chain V r
21	44	100.0	98	2 S29545	Ig heavy chain V r
22	44	100.0	98	2 S26927	Ig heavy chain V r
23	44	100.0	98	2 S26932	Ig heavy chain V r
24	44	100.0	98	2 S26891	Ig heavy chain V r
25	44	100.0	98	2 S26894	Ig heavy chain V r
26	44	100.0	98	2 S26889	Ig heavy chain V r
27	44	100.0	98	2 S26933	Ig heavy chain V r
28	44	100.0	98	2 S26934	Ig heavy chain V-I
29	44	100.0	100	2 PI0122	Ig heavy chain V-I

30	44	100.0	100	2 S69896	Ig heavy chain V r
31	44	100.0	100	2 S26925	Ig heavy chain V r
32	44	100.0	100	2 S26926	Ig heavy chain V r
33	44	100.0	101	2 PU0003	Ig heavy chain V r
34	44	100.0	101	2 PU0002	Ig heavy chain V r
35	44	100.0	105	2 S38488	Ig heavy chain - h
36	44	100.0	105	2 PI0255	Ig heavy chain V r
37	44	100.0	106	2 S24521	Ig heavy chain V r
38	44	100.0	111	1 MHMS76	Ig heavy chain V-I
39	44	100.0	111	2 PH1028	Ig heavy chain V r
40	44	100.0	111	2 S69911	Ig V-D-J region (R
41	44	100.0	112	2 S13693	Ig heavy chain V r
42	44	100.0	112	2 S13690	Ig heavy chain V r
43	44	100.0	113	1 AVMSAB	Ig heavy chain V-I
44	44	100.0	113	1 AVMS61	Ig heavy chain V-I
45	44	100.0	113	1 AVMS09	Ig heavy chain V-I

ALIGNMENTS

RESULT 1
D31485
Ig heavy chain V regions (3-24 and 5-14) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 30-May-1997
C:Accession: D31485
R:Bedzyk, W.D.; Johnson, L.S.; Riordan, G.S.; Voss Jr., E.W.
J. Biol. Chem. 264, 1565-1569, 1989
A:Title: Comparison of variable region primary structures within an anti-fluorescein
A:Reference number: A31485; MUID:89109167
A:Accession: D31485
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-36 <BED>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 44; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGGLVQPG 8
Db 8 GGGLVQPG 15
RESULT 2
S33402
Ig heavy chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 17-Jul-1998
C:Accession: S33402; S36385
R:Kettlborough, C.A.; Saldanha, J.; Ansell, K.H.; Bendig, M.M.
Eur. J. Immunol. 23, 206-211, 1993
A:Title: Optimization of primers for cloning libraries of mouse immunoglobulin genes
A:Reference number: S33391; MUID:9312092
A:Accession: S33402
A:Molecule type: mRNA
A:Residues: 1-38 <KET>
A:Cross-References: EMBL:X73019
A:Experimental source: strain BALB/c
R:Ansell, K.H.
submitted to the EMBL Data Library, April 1993
A:Reference number: S36376
A:Accession: S36385
A:Molecule type: mRNA
A:Residues: 1-30 <ANS>
A:Cross-References: EMBL:X73019
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-38/Domain: immunoglobulin homology (fragment) <IMM>

Query Match
Best Local Similarity 100.0%; Score 44; DB 2; Length 38;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGIVQPG 8
Db 6 GGGIVQPG 13

RESULT 3

A27606
Ig heavy chain V- α region (p26.9p2) - rabbit (fragments)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 08-Mar-1989 #sequence_revision 30-Jun-1991 #text_change 16-Aug-1996
C:Accession: A27606
R:Currier, S.-J.; Gallarda, J.L.; Knight, K.L.
J. Immunol. 140, 1651-1659, 1988
A:Title: Partial molecular genetic map of the rabbit V-H chromosomal region.
A:Reference number: A27606; MUID:88154464
A:Accession: A27606
A:Molecule type: DNA
A:Residues: 1-59 <CUR>
A:Cross-references: GB:M19706
A:Note: the authors translated the codon TCC for residue 28 as Asp
C:Genetics: 30/3
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match
Best Local Similarity 100.0%; Score 44; DB 2; Length 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGIVQPG 8
Db 8 GGGIVQPG 15

RESULT 4

S36381
Ig heavy chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 23-Jul-1999
C:Accession: S36381; S33396
R:Ansell, K.H.
submitted to the EMBL Data Library, April 1993
A:Reference number: S36376
A:Accession: S36381
A:Molecule type: mRNA
A:Residues: 1-59 <ANS>
A:Cross-references: EMBL:X73012; NID:g295879; PIDN:CA51498.1; PID:g939941
Eur. J. Immunol. 23, 206-211, 1993
R:Kettlborough, C.A.; Saldanha, J.; Ansell, K.H.; Bendig, M.M.
A:Title: Optimization of primers for cloning libraries of mouse immunoglobulin genes using
A:Reference number: S33391; MUID:93122092
A:Accession: S33396
A:Molecule type: mRNA
A:Residues: 3-43 <KET>
A:Cross-references: EMBL:X73012
A:Experimental source: strain BALB/c
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:14-59/Domain: immunoglobulin homology (fragment) <IMM>

Query Match
Best Local Similarity 100.0%; Score 44; DB 2; Length 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 10 GGGIVQPG 17

RESULT 5

S36382
Ig heavy chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 23-Jul-1999
C:Accession: S36382; S33397
R:Ansell, K.H.
submitted to the EMBL Data Library, April 1993
A:Reference number: S36376
A:Accession: S36382
A:Molecule type: mRNA
A:Residues: 1-60 <ANS>
A:Cross-references: EMBL:X73011; NID:g295878; PIDN:CA51497.1; PID:g939940
R:Kettlborough, C.A.; Saldanha, J.; Ansell, K.H.; Bendig, M.M.
Eur. J. Immunol. 23, 206-211, 1993
A:Title: Optimization of primers for cloning libraries of mouse immunoglobulin genes
A:Reference number: S33391; MUID:93122092
A:Accession: S33397
A:Molecule type: mRNA
A:Residues: 1-33 <KET>
A:Cross-references: EMBL:X73011
A:Experimental source: strain BALB/c
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:8-60/Domain: immunoglobulin homology (fragment) <IMM>

Query Match
Best Local Similarity 100.0%; Score 44; DB 2; Length 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGIVQPG 8
Db 1 GGGIVQPG 8

RESULT 6

C34964
Ig heavy chain precursor V-III region (Ab21) - human
C:Species: Homo sapiens (man)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-May-1997
C:Accession: C34964
R:Sanz, I.; Casali, P.; Thomas, J.W.; Notkins, A.L.; Capra, J.D.
J. Immunol. 142, 4054-4061, 1989
A:Title: Nucleotide sequences of eight human natural autoantibody V-H regions reveals
A:Reference number: A92830; MUID:89235232
A:Accession: C34964
A:Molecule type: mRNA
A:Residues: 1-82 <SAN>
A:Cross-references: GB:M26996
A:Note: the sequences shown differs from the authors' translation after residue 56
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match
Best Local Similarity 100.0%; Score 44; DB 2; Length 82;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGIVQPG 8
Db 27 GGGIVQPG 34

RESULT 7

D25913
Ig heavy chain V region (BFL1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 16-Aug-1996

C:Accession: D25913
 R:Lewalter, A.M.; Lin, P.S.; Gearhart, P.J.
 Proc. Natl. Acad. Sci. U.S.A. 84, 2454-2458, 1987
 A:Title: Adult B-cell repertoire is biased toward two heavy-chain variable-region genes
 A:Reference number: A94148; MUID:87175692
 A:Accession: D25913
 A:Molecule type: DNA
 A:Residues: 1-94 <LAW>
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotrimer; Immunoglobulin
 F:11-94/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 44; DB 2; Length 94;
 Best Local Similarity 100.0%; Pred. No. 0.29;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGLVQPG 8
 |||||
 DB 4 GGGLVQPG 11

RESULT 8
 167528
 CD3 antigen homolog - mouse (fragment)
 C:Species: Mus sp. (mouse)
 C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jan-2000
 C:Accession: 167528
 R:Chies, J.A.; Lembezat, M.P.; Freltas, A.A.
 Eur. J. Immunol. 24, 1657-1664, 1994
 A:Title: Entry of B lymphocytes into the persistent cell pool in non-immunized mice is
 A:Reference number: 153392; MUID:94298870
 A:Accession: 167528
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-94 <RES>
 A:Cross-references: GB:S71357; NID:9550040
 C:Genetics:
 A:Gene: Ig VHx24b
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 F:11-94/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 44; DB 2; Length 94;
 Best Local Similarity 100.0%; Pred. No. 0.29;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGLVQPG 8
 |||||
 DB 4 GGGLVQPG 11

RESULT 9
 167527
 CD3 antigen homolog - mouse (fragment)
 C:Species: Mus sp. (mouse)
 C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jan-2000
 C:Accession: 167527
 R:Chies, J.A.; Lembezat, M.P.; Freltas, A.A.
 Eur. J. Immunol. 24, 1657-1664, 1994
 A:Title: Entry of B lymphocytes into the persistent cell pool in non-immunized mice is
 A:Reference number: 153392; MUID:94298870
 A:Accession: 167527
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-95 <RES>
 A:Cross-references: GB:S71351; NID:9550039
 C:Genetics:
 A:Gene: Ig VHx24a
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 F:12-95/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 44; DB 2; Length 95;
 Best Local Similarity 100.0%; Pred. No. 0.29;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGLVQPG 8
 |||||
 DB 5 GGGLVQPG 12

RESULT 10
 PH0872
 Ig heavy chain V region (anti-DNA, III-3R) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 16-Aug-1996
 C:Accession: PH0872
 R:Manheimer-Lory, A.; Katz, J.B.; Pillinger, M.; Ghossein, C.; Smith, A.; Diamond, B.
 J. Exp. Med. 174, 1639-1652, 1991
 A:Title: Molecular characteristics of antibodies bearing an anti-DNA-associated idiot
 A:Reference number: PH0862; MUID:92078875
 A:Accession: PH0872
 A:Molecule type: DNA
 A:Residues: 1-97 <MAN>
 A:Note: a "G" was inserted at the position of between 72nd and 73rd of the DNA sequen
 C:Comment: This antibody is produced by Epstein-Barr virus-transformed B cell that be
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotrimer; Immunoglobulin
 F:14-97/Domain: Immunoglobulin homology <IMM>
 F:30-35/Region: complementarity-determining 1
 F:49-66/Region: complementarity-determining 2

Query Match 100.0%; Score 44; DB 2; Length 97;
 Best Local Similarity 100.0%; Pred. No. 0.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGLVQPG 8
 |||||
 DB 7 GGGLVQPG 14

RESULT 11
 S26885
 Ig heavy chain V region (DP-44) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 23-Jul-1999
 C:Accession: S26885; S36596
 R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
 J. Mol. Biol. 227, 776-798, 1992
 A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups o
 A:Reference number: S26885; MUID:93021117
 A:Accession: S26885
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-97 <TOM1>
 A:Cross-references: EMBL:Z12344
 R:Tomlinson, M.
 submitted to the EMBL Data Library, June 1992
 A:Reference number: S36596
 A:Accession: S36596
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-5, 'Q', '7-12, 'H', '14-97 <TOM2>
 A:Cross-references: EMBL:Z12344; NID:932908; PIDN:CAA78214.1; PID:932909
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotrimer; Immunoglobulin
 F:15-97/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 44; DB 2; Length 97;
 Best Local Similarity 100.0%; Pred. No. 0.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGLVQPG 8

Db 8 GGGVQPG 15

RESULT 12

S26886

Ig heavy chain V region (DP-45) - human

C:Species: Homo sapiens (man)

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999

C:Accession: S26886

R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.

J. Mol. Biol. 227, 776-798, 1992

A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of V

A:Reference number: S26885; MUID:93021117

A:Accession: S26886

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-97 <TOM>

A:Cross-references: EMBL:212345; NID:932910; PIDN:CAA78215.1; PID:932911

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 44; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQPG 8
Db 8 GGGVQPG 15

RESULT 13

S26890

Ig heavy chain V region (DP-48) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999

C:Accession: S26890

R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.

J. Mol. Biol. 227, 776-798, 1992

A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of V

A:Reference number: S26885; MUID:93021117

A:Accession: S26890

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-97 <TOM>

A:Cross-references: EMBL:212348; NID:932916; PIDN:CAA78218.1; PID:932917

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 44; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQPG 8
Db 8 GGGVQPG 15

RESULT 14

S46462

Ig heavy chain V region (VAC-5) - human

C:Species: Homo sapiens (man)

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999

C:Accession: S46462

R:Cook, G.P.; Tomlinson, I.M.; Walter, G.; Riethman, H.; Carter, N.P.; Buluwela, L.; Win

Nature Genet. 7, 162-168, 1994

A:Title: A map of the human immunoglobulin V(H) locus completed by analysis of the telom

A:Reference number: S46460; MUID:95004581

A:Accession: S46462

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-97 <COO>
A:Cross-references: EMBL:227504; NID:9505430; PIDN:CAA81824.1; PID:9505431
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 44; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQPG 8
Db 8 GGGVQPG 15

RESULT 15

S26895

Ig heavy chain V region (DP-52) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999

C:Accession: S26895

R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.

J. Mol. Biol. 227, 776-798, 1992

A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups o

A:Reference number: S26885; MUID:93021117

A:Accession: S26895

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-97 <TOM>

A:Cross-references: EMBL:212352; NID:932926; PIDN:CAA78222.1; PID:932927

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 44; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQPG 8
Db 8 GGGVQPG 15

Search completed: June 13, 2001, 14:23:10
Job time: 741 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:30:36 ; Search time 51.57 Seconds
(without alignments)
5.314 Million cell updates/sec

Title: PCT-US01-05825A-23

Perfect score: 44
Sequence: 1 GGGVQPG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	44	100.0	111 1 HV35_MOUSE	P01804 mus musculu
2	44	100.0	113 1 HV27_MOUSE	P01796 mus musculu
3	44	100.0	113 1 HV28_MOUSE	P01797 mus musculu
4	44	100.0	113 1 HV28_MOUSE	P01798 mus musculu
5	44	100.0	113 1 HV30_MOUSE	P01799 mus musculu
6	44	100.0	113 1 HV31_MOUSE	P01800 mus musculu
7	44	100.0	113 1 HV34_MOUSE	P01803 mus musculu
8	44	100.0	115 1 HV32_MOUSE	P01801 mus musculu
9	44	100.0	115 1 HV33_MOUSE	P01802 mus musculu
10	44	100.0	115 1 HV3D_HUMAN	P01765 homo sapien
11	44	100.0	116 1 HV36_MOUSE	P01806 mus musculu
12	44	100.0	116 1 HV3R_HUMAN	P01779 homo sapien
13	44	100.0	117 1 HV17_MOUSE	P01786 mus musculu
14	44	100.0	117 1 HV3C_HUMAN	P01764 homo sapien
15	44	100.0	117 1 HV3O_HUMAN	P01776 homo sapien
16	44	100.0	117 1 HV41_MOUSE	P01811 mus musculu
17	44	100.0	117 1 HV54_MOUSE	P18525 mus musculu
18	44	100.0	118 1 HV39_MOUSE	P01809 mus musculu
19	44	100.0	118 1 HV3V_HUMAN	P80419 homo sapien
20	44	100.0	119 1 HV37_MOUSE	P01807 mus musculu
21	44	100.0	119 1 HV38_MOUSE	P01808 mus musculu
22	44	100.0	119 1 HV3K_HUMAN	P01774 homo sapien
23	44	100.0	119 1 HV3N_HUMAN	P01775 homo sapien
24	44	100.0	119 1 HV3P_HUMAN	P01777 homo sapien
25	44	100.0	119 1 HV40_MOUSE	P01810 mus musculu
26	44	100.0	120 1 HV3E_HUMAN	P01766 homo sapien
27	44	100.0	122 1 HV2O_MOUSE	P01789 mus musculu
28	44	100.0	122 1 HV2I_MOUSE	P01790 mus musculu
29	44	100.0	123 1 HV18_MOUSE	P01787 mus musculu
30	44	100.0	123 1 HV19_MOUSE	P01788 mus musculu
31	44	100.0	123 1 HV22_MOUSE	P01791 mus musculu
32	44	100.0	123 1 HV23_MOUSE	P01792 mus musculu
33	44	100.0	123 1 HV24_MOUSE	P01793 mus musculu

ALIGNMENTS

34	44	100.0	123 1 HV25_MOUSE	P01794 mus musculu
35	44	100.0	136 1 HV16_MOUSE	P01783 mus musculu
36	44	100.0	142 1 HV01_RAT	P01805 ratuus norv
37	44	100.0	144 1 HV26_MOUSE	P01795 mus musculu
38	43	97.7	115 1 HV3F_HUMAN	P01767 homo sapien
39	43	97.7	116 1 HV05_CARAU	P19181 carassius a
40	41	93.2	114 1 HV3B_HUMAN	P01763 homo sapien
41	41	93.2	119 1 HV3I_HUMAN	P01770 homo sapien
42	41	93.2	121 1 HV3J_HUMAN	P01771 homo sapien
43	41	93.2	126 1 HV3K_HUMAN	P01772 homo sapien
44	40	90.9	97 1 HV56_MOUSE	P18527 mus musculu
45	40	90.9	98 1 HV57_MOUSE	P18528 mus musculu

RESULT 1
HV35_MOUSE STANDARD; PRT: 111 AA.
AC P01804;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION HPC76 (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81013937; PubMed=6251474;
RA Bernard O., Gough N.M.;
RT "Nucleotide sequence of immunoglobulin heavy chain joining segments
between translocated VH and mu constant regions.";
RL Proc. Natl. Acad. Sci. U.S.A. 77:3630-3634(1980).
CC -!- MISCELLANEOUS: THE SEQUENCE OF THE FIRST 197 RESIDUES OF THE C
REGION WAS ALSO DETERMINED & DIFFERS IN ONLY 3 POSITIONS FROM THE
CC CORRESPONDING PORTION OF THE MOUSE MOPC 104E MU CHAIN.
DR PIR: A02074; MHMS76.
DR InterPro: IPR003006; .
DR Pfam: PF00047; Ig: 1.
KW Immunoglobulin V region.
FT NON_TER 1 1
FT NON_TER 111 111
SO SEQUENCE 111 AA; 12304 MW; 0EDE98EC7348056A CRC64;

Query Match 100.0%; Score 44; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGVQPG 8
Db 3 GGGVQPG 10

RESULT 2
HV27_MOUSE STANDARD; PRT: 113 AA.
AC P01756;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION A4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=78158406; PubMed=417344;
RA Vrana M., Rudikoff S., Potter M.;

RT "Sequence variation among heavy chains from inulin-binding myeloma proteins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 75:1957-1961(1978).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
 CC BINDS INULIN.
 DR PIR: A93818; AVMSAR.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 19; 1.
 FT Immunoglobulin V region.
 FT DISULFID 22 98
 FT NON_TER 113 113 BY SIMILARITY.
 SQ SEQUENCE 113 AA; 12675 MW; 76658C121C598285 CRC64;

Query Match 100.0%; Score 44; DB 1; Length 113;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGLVOPG 8
 DB 8 GGGLVOPG 15

RESULT 3
 HV28_MOUSE STANDARD; PRT; 113 AA.
 AC P01797;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION U61.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=78158406; PubMed=417344;
 RA Vrana M., Rudikoff S., Potter M.;
 RT "Sequence variation among heavy chains from inulin-binding myeloma
 RT proteins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 75:1957-1961(1978).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
 CC BINDS INULIN.
 DR PIR: B93818; AVMS61.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region.
 FT DISULFID 22 98
 FT NON_TER 113 113 BY SIMILARITY.
 SQ SEQUENCE 113 AA; 12671 MW; 76634C121C598285 CRC64;

Query Match 100.0%; Score 44; DB 1; Length 113;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGLVOPG 8
 DB 8 GGGLVOPG 15

RESULT 4
 HV29_MOUSE STANDARD; PRT; 113 AA.
 AC P01798;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION E109.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=78158406; PubMed=417344;
 RA Vrana M., Rudikoff S., Potter M.;
 RT "Sequence variation among heavy chains from inulin-binding myeloma
 RT proteins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 75:1957-1961(1978).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
 CC BINDS INULIN.
 DR PIR: C93818; AVMS09.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region.
 FT DISULFID 22 98
 FT NON_TER 113 113 BY SIMILARITY.
 SQ SEQUENCE 113 AA; 12647 MW; EE50F2F20EDB129B CRC64;

Query Match 100.0%; Score 44; DB 1; Length 113;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGLVOPG 8
 DB 8 GGGLVOPG 15

RESULT 5
 HV30_MOUSE STANDARD; PRT; 113 AA.
 AC P01799;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION ABE-47N.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=77134726; PubMed=402936;
 RA Vrana M., Rudikoff S., Potter M.;
 RT "Heavy-chain variable-region sequence from an inulin-binding myeloma
 RT protein.";
 RL Biochemistry 16:1170-1175(1977).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
 CC BINDS INULIN.
 DR PIR: A90400; AVMS87.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region.
 FT DISULFID 22 98
 FT NON_TER 113 113 BY SIMILARITY.
 SQ SEQUENCE 113 AA; 12675 MW; 76658C16C779845E CRC64;

Query Match 100.0%; Score 44; DB 1; Length 113;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGLVOPG 8
 DB 8 GGGLVOPG 15

RESULT 6
 HV31_MOUSE STANDARD; PRT; 113 AA.
 AC P01800;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION T957.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=81216632; PubMed=6787122;
RA Rudkoff S., Potter M.;
RT "Immunoglobulin heavy chains from anti-inulin myeloma proteins:
RT evidence for a new heavy chain joining segment.";
RL J. Immunol. 127:191-194(1981).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS INULIN.
DR PIR: A92810; AVMS57.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT DISULFID 22 98
FT NON_TER 113 113 BY SIMILARITY.
SQ SEQUENCE 113 AA; 12732 MW; 26618F626B59859E CRC64;

Query Match 100.0%; Score 44; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGLVQPG 8
DB 8 GGGLVQPG 15

RESULT 7
HV34_MOUSE
ID HV34_MOUSE STANDARD; PRT; 113 AA.
AC P01803;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V REGION AMPCL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=81216632; PubMed=6787122;
RA Rudkoff S., Potter M.;
RT "Immunoglobulin heavy chains from anti-inulin myeloma proteins:
RT evidence for a new heavy chain joining segment.";
RL J. Immunol. 127:191-194(1981).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS INULIN.
DR PIR: A02073; HVMSAM.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT DISULFID 22 98 BY SIMILARITY.
FT NON_TER 113 113
SQ SEQUENCE 113 AA; 12691 MW; 7A6D906AAA966E9E CRC64;

Query Match 100.0%; Score 44; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGLVQPG 8
DB 8 GGGLVQPG 15

RESULT 8

HV32_MOUSE
ID HV32_MOUSE STANDARD; PRT; 115 AA.
AC P01801;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION J606.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=82099361; PubMed=6798111;
RA Johnson N., Stankard J., Paul L., Hood L.;
RT "The complete V domain amino acid sequences of two myeloma inulin-
RT binding proteins.";
RL J. Immunol. 128:302-307(1982).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS INULIN.
DR PIR: G92811; AVMS06.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT DISULFID 22 98 BY SIMILARITY.
FT NON_TER 115 115
SQ SEQUENCE 115 AA; 12810 MW; B67AD6638A12A5F CRC64;

Query Match 100.0%; Score 44; DB 1; Length 115;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGLVQPG 8
DB 8 GGGLVQPG 15

RESULT 9
HV33_MOUSE
ID HV33_MOUSE STANDARD; PRT; 115 AA.
AC P01802;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION W3082.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=82099361; PubMed=6798111;
RA Johnson N., Stankard J., Paul L., Hood L.;
RT "The complete V domain amino acid sequences of two myeloma inulin-
RT binding proteins.";
RL J. Immunol. 128:302-307(1982).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS INULIN.
DR PIR: D92811; AVMS82.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT DISULFID 22 98 BY SIMILARITY.
FT NON_TER 115 115
SQ SEQUENCE 115 AA; 12887 MW; 9B4517648C121C5A CRC64;

Query Match 100.0%; Score 44; DB 1; Length 115;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGLVQPG 8

```

Db      8 GGGLVQPG 15

RESULT 10
HV3D_HUMAN STANDARD; PRT; 115 AA.
ID HV3D_HUMAN
AC P01765;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION TITL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=78005528; PubMed=409716;
RA Wang A.-C., Wang T.Y., Fudenberg H.H.;
RT "Immunoglobulin structure and genetics. Identity between variable
RT regions of a mu and a gamma2 chain.";
RL J. Biol. Chem. 252:7192-7199(1977).
CC -1- MISCELLANEOUS: THE SEQUENCES OF THE V REGIONS OF THE HEAVY CHAINS
CC OF IGM AND IGG2 ISOLATED FROM A SINGLE PATIENT WITH BICLONAL
CC GAMMOPATHY ARE IDENTICAL. THEIR LIGHT CHAINS ARE APPARENTLY ALSO
CC IDENTICAL.
DR PIR: A02048; H3HUTL.
DR InterPro: IPR003006;
DR Pfam: PF00047; Ig: 1.
DR Immunoglobulin V region.
FT NON_TER 115
SQ SEQUENCE 115 AA; 12356 MW; 4DCC67D179F62326 CRC64;

Query Match 100.0%; Score 44; DB 1; Length 115;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGLVQPG 8
Db 8 GGGLVQPG 15

RESULT 11
HV36_MOUSE STANDARD; PRT; 116 AA.
ID HV36_MOUSE
AC P01806;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V REGION 441 PRECURSOR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=82059449; PubMed=6795591;
RA Olio R., Aufray C., Sikorav J.-L., Rougeon F.;
RT "Mouse heavy chain variable regions: nucleotide sequence of a
RT germ-line VH gene segment.";
RL Nucleic Acids Res. 9:4099-4109(1981).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: V00774; CAA24152.1;

```

```

DR PIR: A02076; HWSA4.
DR HSSP: P01810; 2FE8.
DR InterPro: IPR003006;
DR Pfam: PF00047; Ig: 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1
FT CHAIN 19 IG HEAVY CHAIN V REGION 441.
FT NON_TER 116
SQ SEQUENCE 116 AA; 12911 MW; 6B4FE3EF9A95D9F3 CRC64;

Query Match 100.0%; Score 44; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGLVQPG 8
Db 26 GGGLVQPG 33

RESULT 12
HV3R_HUMAN STANDARD; PRT; 116 AA.
ID HV3R_HUMAN
AC P01779;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION TUR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=74142702; PubMed=4522793;
RA Capra J.D., Kehoe J.M.;
RT "Variable region sequences of five human immunoglobulin heavy chains
RT of the VH3 subgroup: definitive identification of four heavy chain
RT hypervariable regions.";
RL Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGAI MYELOMA
CC PROTEIN.
CC
CC PIR: A02062; A1HUTU.
DR InterPro: IPR003006;
DR Pfam: PF00047; Ig: 1.
DR Immunoglobulin V region.
FT NON_TER 116
SQ SEQUENCE 116 AA; 12431 MW; EB705F553A963F0C CRC64;

Query Match 100.0%; Score 44; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGLVQPG 8
Db 8 GGGLVQPG 15

RESULT 13
HV17_MOUSE STANDARD; PRT; 117 AA.
ID HV17_MOUSE
AC P01786;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V REGION MOPC 47A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.

```

```

RX MEDLINE=80049769; PubMed=115869;
RA Robinson E.A., Appella E.;
RT "Antino acid sequence of a mouse myeloma immunoglobulin heavy chain
RT (MOPC 47 A) with a 100-residue deletion."
RL J. Biol. Chem. 254:11418-11430(1979).
CC -I- MISCELLANEOUS: THIS ALPHA CHAIN WAS ISOLATED FROM A MYELOMA
CC PROTEIN THAT CONTAINS ONE LIGHT AND ONE HEAVY CHAIN PER MOLECULE,
CC LINKED BY A DISULFIDE BOND. IN CONTRAST, NORMAL MOUSE IGA
CC MOLECULES CONTAIN TWO LIGHT AND TWO HEAVY CHAINS AND LACK A
CC LIGHT-HEAVY CHAIN DISULFIDE BOND.
DR PIR: A02069; AIMS47.
DR HSSP: P01789; 2MCP.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 1g: 1.
DR Immunoglobulin V region.
KW NON_TER 117 117
SQ SEQUENCE 117 AA; 12975 MW; 0C74BB8B154BDF4 CRC64;

Query Match 100.0%; Score 44; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGLVQPG 8
   |||||
Db 8 GGGLVQPG 15

RESULT 14
HV3C_HUMAN STANDARD; PRT; 117 AA.
AC P01764;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION VH26 PRECURSOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.;
RX MEDLINE=81101090; PubMed=6450418;
RA Matthysens G., Rablitz T.H.;
RT "Structure and multiplicity of genes for the human immunoglobulin
RT heavy chain variable region."
RL Proc. Natl. Acad. Sci. U.S.A. 77:6561-6565(1980).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00236; AAAS3516.1; -.
DR EMBL: M35415; AAAS8735.1; -.
DR PIR: A02047; H3H026.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 1g: 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V-III REGION VH26.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12582 MW; EB26733F1A3CB0F1 CRC64;

Query Match 100.0%; Score 44; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGLVQPG 8

```

```

Db 27 GGGLVQPG 34
   |||||
RESULT 15
HV3C_HUMAN STANDARD; PRT; 117 AA.
AC P01776;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION WAS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=74142702; PubMed=4522793;
RA Capra J.D., Kenne J.M.;
RT "Variable region sequences of five human immunoglobulin heavy chains
RT of the VH3 subgroup: definitive identification of four heavy chain
RT hypervariable regions."
RL Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA
CC PROTEIN.
DR PIR: A02059; G1H0WS.
DR HSSP: P01772; 2IG2.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 1g: 1.
KW Immunoglobulin V region.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 13091 MW; 201DEF0E1E5309BF CRC64;

Query Match 100.0%; Score 44; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGLVQPG 8
   |||||
Db 8 GGGLVQPG 15

Search completed: June 13, 2001, 14:30:36
Job time: 526 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:29:41 ; Search time 150.43 Seconds
(without alignments)
6.233 Million cell updates/sec

Title: PCT-US01-05825A-23

Perfect score: 44

Sequence: 1 GGGLVQPG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_15:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mbc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_protist:*
- 12: sp_unclassified:*
- 13: sp_vertebrate:*
- 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	15	11	09QV16
2	44	100.0	95	4	09UL86
3	44	100.0	118	4	09UL91
4	44	100.0	118	4	09UL72
5	41	93.2	113	4	09UL90
6	41	93.2	116	4	09UL93
7	41	93.2	121	4	09UL71
8	41	93.2	122	4	09UL84
9	41	93.2	147	4	09Y509
10	41	93.2	168	2	09S2V7
11	40	90.9	131	4	09UL88
12	40	90.9	298	11	09QVFO
13	40	90.9	437	11	09RLA4
14	38	86.4	731	5	09Y2K7
15	36	81.8	249	2	084994
16	36	81.8	354	5	09X2Y2
17	36	81.8	382	2	09WMT2
18	36	81.8	382	2	09S640
19	36	81.8	383	2	086045

20	36	81.8	383	2	09X3A8	09X3A8 burkholderi
21	36	81.8	384	2	068143	068143 burkholderi
22	36	81.8	579	4	09UE17	09ue17 homo sapien
23	35	79.5	16	4	09UC53	09uc53 homo sapien
24	35	79.5	197	5	09N3C9	09n3c9 caenorhabdi
25	35	79.5	238	2	09PJT5	09pjt5 chlamydia m
26	35	79.5	268	2	09J0I2	09j0i2 chlamydia p
27	35	79.5	268	4	09S616	09s616 homo sapien
28	35	79.5	323	13	073907	073907 gallus gall
29	35	79.5	855	5	09N3U0	09n3u0 caenorhabdi
30	34	77.3	225	2	09RXW3	09rxw3 deinococcus
31	34	77.3	254	2	09R9B1	09r9b1 burkholderi
32	34	77.3	336	5	09VB05	09vb05 drosophila
33	34	77.3	387	2	09WMT1	09wmt1 burkholderi
34	34	77.3	387	2	09S639	09s639 burkholderi
35	34	77.3	387	2	09S638	09s638 burkholderi
36	34	77.3	387	2	09ROE9	09roe9 burkholderi
37	34	77.3	388	2	09ZFK3	09zfk3 burkholderi
38	34	77.3	388	2	P70944	P70944 burkholderi
39	34	77.3	421	4	09NP21	09np21 homo sapien
40	34	77.3	445	4	09NP18	09np18 homo sapien
41	34	77.3	446	4	09NP19	09np19 homo sapien
42	34	77.3	451	4	09NP20	09np20 homo sapien
43	34	77.3	455	4	09UP88	09up88 homo sapien
44	34	77.3	456	4	09Y215	09y215 homo sapien
45	34	77.3	481	2	09KKM3	09kkm3 vitrio chol

ALIGNMENTS

RESULT 1

09QV16 ID 09QV16 PRELIMINARY: PRT: 15 AA.

AC 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE PROLACTIN-BINDING PROTEIN (FRAGMENT).

OS Rattus sp.

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

OX NCBI_TaxID=10118;

RN [1]

RP SEQUENCE.

RX MEDLINE=95094032; PubMed=8000909;

RA Cohen H., Cohen O., Gagnon J.;

RT "Serum prolactin-binding protein (PRL-BP) of human and rat are identified as IgG.";

RL C. R. Acad. Sci., III, Sci. Vie 317:293-298(1994).

DR HSSP; P01789; IMCP.

SQ SEQUENCE 15 AA; 1469 MW; 35ED2512FF3FA369 CRC64;

Query Match 100.0%; Score 44; DB 11; Length 15;
Best local Similarity 100.0%; Pred. No. 0.094; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQPG 8
DB 8 GGGVQPG 15

RESULT 2

09UL86 ID 09UL86 PRELIMINARY: PRT: 95 AA.

AC 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE IMMUNOGLOBULIN HEAVY CHAIN (FRAGMENT).

VH.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OX Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN NCBI_TaxID=9606;
 RP [1]
 RP SEQUENCE FROM N.A.
 RA Tange Y., Kayano H.;
 RT "Human VH gene sequence."
 RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AB035268; BAA87067.1; -
 DR HSSP; P01772; 2FB4.
 DR INTERPRO: IPR003006; -
 DR PFAM: PF00047; 19; 1.
 FT NON_TER 1 1
 FT NON_TER 95 95
 SQ SEQUENCE 95 AA; 10527 MW; 90A8CD16D22574A CRC64;

Query Match 100.0%; Score 44; DB 4; Length 95;
 Best Local Similarity 100.0%; Pred. No. 0.66;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVOPG 8
 |||||
 DB 7 GGGVOPG 14

RESULT 3
 Q9UL91 PRELIMINARY; PRT; 118 AA.
 ID Q9UL91;
 AC Q9UL91;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 OS MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OX Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN NCBI_TaxID=9606;
 RP [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-98277139; PubMed-9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus."
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL; AF035023; AAD56259.1; -
 DR INTERPRO: IPR003006; -
 DR PFAM: PF00047; 19; 1.
 FT NON_TER 1 1
 FT NON_TER 118 118
 SQ SEQUENCE 118 AA; 12843 MW; D0633949F2AC149D CRC64;

Query Match 100.0%; Score 44; DB 4; Length 118;
 Best Local Similarity 100.0%; Pred. No. 0.86;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVOPG 8
 |||||
 DB 8 GGGVOPG 15

RESULT 4
 Q9UL72 PRELIMINARY; PRT; 118 AA.
 ID Q9UL72;
 AC Q9UL72;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 OS MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
 OC Homo sapiens (human).
 DE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-98277139; PubMed-9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus."
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL; AF035042; AAD56278.1; -
 DR HSSP; P01772; 2FB4.
 DR INTERPRO: IPR003006; -
 DR PFAM: PF00047; 19; 1.
 FT NON_TER 1 1
 FT NON_TER 118 118
 SQ SEQUENCE 118 AA; 12872 MW; B4D1A5944B2D5CCA CRC64;

Query Match 100.0%; Score 44; DB 4; Length 118;
 Best Local Similarity 100.0%; Pred. No. 0.86;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVOPG 8
 |||||
 DB 8 GGGVOPG 15

RESULT 5
 Q9UL90 PRELIMINARY; PRT; 113 AA.
 ID Q9UL90;
 AC Q9UL90;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 OS MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
 OC Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OX Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN NCBI_TaxID=9606;
 RP [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-98277139; PubMed-9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus."
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL; AF035024; AAD56260.1; -
 DR INTERPRO: IPR003006; -
 DR PFAM: PF00047; 19; 1.
 FT NON_TER 1 1
 FT NON_TER 113 113
 SQ SEQUENCE 113 AA; 12437 MW; ED57FDD19086D07F CRC64;

Query Match 93.2%; Score 41; DB 4; Length 113;
 Best Local Similarity 87.5%; Pred. No. 2.9;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVOPG 8
 |||||
 DB 8 GGGVOPG 15

RESULT 6
 Q9UL93 PRELIMINARY; PRT; 116 AA.
 ID Q9UL93;
 AC Q9UL93;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 OS MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
 DE

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9827139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalls N.N., Berney S.M.,
Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035021; AAD56257.1; -.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; 19; 1.
FT NON_TER 1 1
FT 116 116
SO SEQUENCE 116 AA; 12434 MW; 0DA0348154DD061 CRC64;

Query Match 93.2%; Score 41; DB 4; Length 116;
Best Local Similarity 87.5%; Pred. No. 2.9;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVOPG 8
Db 7 GGGVOPG 14

RESULT 7
O9UL71 PRELIMINARY; PRT; 121 AA.
AC O9UL71;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9827139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalls N.N., Berney S.M.,
Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035043; AAD56279.1; -.
DR HSSP: P01772; 2FB4.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; 19; 1.
FT NON_TER 1 1
FT 121 121
SO SEQUENCE 121 AA; 13154 MW; 2F045CCEA5D50736 CRC64;

Query Match 93.2%; Score 41; DB 4; Length 121;
Best Local Similarity 87.5%; Pred. No. 3.1;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVOPG 8
Db 8 GGGVOPG 15

RESULT 8
O9UL84 PRELIMINARY; PRT; 122 AA.
AC O9UL84;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9827139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalls N.N., Berney S.M.,
Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035030; AAD56266.1; -.
DR HSSP: P01772; 2FB4.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; 19; 1.
FT NON_TER 1 1
FT 122 122
SO SEQUENCE 122 AA; 13579 MW; 36054D41366545B8 CRC64;

Query Match 93.2%; Score 41; DB 4; Length 122;
Best Local Similarity 87.5%; Pred. No. 3.1;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVOPG 8
Db 8 GGGVOPG 15

RESULT 9
O9Y509 PRELIMINARY; PRT; 147 AA.
AC O9Y509;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DE VHS3 PROTEIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96071149; PubMed=7475288;
RA Cao J., Vesic R.A., Rettig M.B., Hong C.H., Kim A., Lee J.C.,
Lichtenstein A.K., Berenson J.R.;
RT "A CD10-positive subset of malignant cells is identified in multiple
myeloma using PCR with patient-specific immunoglobulin gene primers."
RL Leukemia 9:1948-1953(1995).
DR EMBL: S80860; AAD14339.1; -.
DR HSSP: P01772; 2FB4.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; 19; 1.
FT NON_TER 1 1
FT 147 147
SO SEQUENCE 147 AA; 15768 MW; 8489FCANA7BC925C CRC64;

Query Match 93.2%; Score 41; DB 4; Length 147;
Best Local Similarity 87.5%; Pred. No. 3.8;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVOPG 8
Db 8 GGGVOPG 15

RESULT 10
O9SZV7 PRELIMINARY; PRT; 168 AA.
ID O9SZV7

AC 09S2V7;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 GN HYPOTHETICAL 17.2 KDA PROTEIN.
 OS Streptomyces coelicolor.
 OC Bacteria; Filicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Redenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 RL M01. Microbiol. 21:77-96(1996).
 DR EMBL: AL109663; CAB52008.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 168 AA; 17181 MW; D950A20F2267097B CRC64;

Query Match 93.2%; Score 41; DB 2; Length 168;
 Best Local Similarity 87.5%; Pred. No. 4.4;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGLVQPG 8
 Db 132 GGGVLQPG 139
 |||||

RESULT 11
 Q9UL88 PRELIMINARY; PRT; 131 AA.
 AC Q9UL88;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9827139; PubMed=9614934;
 RA Wu X., Liu B., Van der Werwe P.L., Kalis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 RT fetus".
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL: AF035026; AAD56262.1; -.
 DR INTERPRO: IPR003006; -.
 DR PIRAM; PF00047; 19; 1.
 FT NON_TER 131 131
 FT 1
 SQ SEQUENCE 131 AA; 14142 MW; 96E7D668E375DEA0 CRC64;

Query Match 90.9%; Score 40; DB 4; Length 131;
 Best Local Similarity 87.5%; Pred. No. 5.1;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGLVQPG 8
 Db 8 GGGLVQPG 15
 |||||

RESULT 12
 Q90YF0 PRELIMINARY; PRT; 298 AA.
 AC Q90YF0;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 GN CN 8 SCFV.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE=SPLEEN;
 RA Shinozaki N., Demura T., Fukuda H.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE=SPLEEN;
 RA Shinozaki N., Demura T., Fukuda H.;
 RT "Isolation of a novel type of vascular cell wall-specific monoclonal
 RT antibody recognizing a cell polarity using a phage display subtraction
 RT method."
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB036341; BAA88633.1; -.
 DR HSSP: P01607; 1RET
 DR INTERPRO: IPR003006; -.
 DR PIRAM; PF00047; 19; 2.
 SQ SEQUENCE 298 AA; 31867 MW; E0F96B8A17004317 CRC64;

Query Match 90.9%; Score 40; DB 11; Length 298;
 Best Local Similarity 87.5%; Pred. No. 12;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGLVQPG 8
 Db 47 GGGLVQPG 54
 |||||

RESULT 13
 Q9RI44 PRELIMINARY; PRT; 437 AA.
 AC Q9RI44;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE GAMMA1 HEAVY CHAIN OF MAB7 (FRAGMENT).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
 RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
 RT antibody (Mab 7, its light and heavy chains) and construction of a
 RT single chain antibody (scFv)."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF152372; AAD40243.1; -.
 DR HSSP: P01842; 7FAB.
 DR INTERPRO: IPR003006; -.
 DR PIRAM; PF00047; 19; 4.
 DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
 FT NON_TER 1 1

FT NON_TER 437 437
SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D697C CRC64;

Query Match
Best Local Similarity 90.9%; Score 40; DB 11; Length 437;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVOPG 8
| | | | |
DB 7 GGGVOPG 14

RESULT 14
O9VZK7 PRELIMINARY; PRT: 731 AA.
AC O9VZK7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE CG14982 PROTEIN.
CN CG14982.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champ M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
RA Burlile K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,
RA Jaisli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Klimek B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclio J.M.,
RA Palazzolo M., Plittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003479; AAF47814.1; -;
DR FLYBASE: FBgn0035477; CG14982.
SQ SEQUENCE 731 AA; 80099 MW; D346AFC51B7F0CE1 CRC64;

Query Match
Best Local Similarity 86.4%; Score 38; DB 5; Length 731;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGVOPG 8
| | | | |
DB 576 GGGVOPG 583

RESULT 15
O84994 PRELIMINARY; PRT: 249 AA.
AC O84994;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE FLAGELLIN (FRAGMENT)
OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC Burkholderia.
NX NCBI_TaxID=28450;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-E956;
RX MEDLINE=99093016; PubMed=9877189;
RA Winstanley C., Hales B.A., Corkill J.E., Gallagher M.J., Hart C.A.;
RT "Flagellin gene variation between clinical and environmental isolates
RT of Burkholderia pseudomallei contrasts with the invariance among
RT clinical isolates.";
RL J. Med. Microbiol. 47:689-694(1998).
DR EMBL: AF030241; AAC27443.1; -;
FT NON_TER 1 1
FT NON_TER 249 249
SQ SEQUENCE 249 AA; 24915 MW; EFF4C8CA527174C3 CRC64;

Query Match
Best Local Similarity 81.8%; Score 36; DB 2; Length 249;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGVOPG 8
| | | | |
DB 74 GGGVOPG 81

Search completed: June 13, 2001, 14:29:42
Job time: 543 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: June 13, 2001, 14:27:06 ; Search time 78.71 Seconds
(without alignments)
1.953 Million cell updates/sec

Title: PCT-US01-05825A-23

Perfect score: 44

Sequence: 1 GCGLVQPG 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Issued Patents, AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/Backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	15	1	US-08-331-398A-66
2	44	100.0	15	2	US-08-652-558-42
3	44	100.0	15	2	US-08-331-397B-66
4	44	100.0	15	2	US-08-758-804A-65
5	44	100.0	18	1	US-08-331-398A-55
6	44	100.0	18	1	US-08-401-908-14
7	44	100.0	18	2	US-08-331-397B-55
8	44	100.0	18	2	US-08-759-804A-54
9	44	100.0	20	2	US-08-859-931A-2
10	44	100.0	20	2	US-08-053-451B-113
11	44	100.0	26	1	US-08-471-780C-80
12	44	100.0	26	1	US-08-467-282B-80
13	44	100.0	26	1	US-08-471-282A-80
14	44	100.0	26	2	US-08-466-710C-80
15	44	100.0	26	3	US-08-468-739C-80
16	44	100.0	29	1	US-08-471-780C-119
17	44	100.0	29	1	US-08-467-282B-119
18	44	100.0	29	2	US-08-471-282A-119
19	44	100.0	29	3	US-08-466-710C-119
20	44	100.0	29	3	US-08-468-739C-119
21	44	100.0	30	1	US-07-988-925-7
22	44	100.0	30	1	US-07-988-925-7
23	44	100.0	30	1	US-08-129-930B-75
24	44	100.0	30	2	US-08-362-780-7
25	44	100.0	30	2	US-08-765-783A-92
26	44	100.0	30	2	US-08-470-139-17
27	44	100.0	67	1	US-08-162-102C-36

28	44	100.0	86	2	US-08-053-451B-126	Sequence 126, App
29	44	100.0	87	1	US-08-497-312-21	Sequence 21, App
30	44	100.0	87	2	US-08-765-783A-105	Sequence 105, App
31	44	100.0	98	1	US-08-211-202-116	Sequence 116, App
32	44	100.0	98	1	US-07-942-245-37	Sequence 37, App
33	44	100.0	98	2	US-08-428-197-48	Sequence 48, App
34	44	100.0	98	2	US-08-665-202-31	Sequence 31, App
35	44	100.0	98	5	PCT-US93-10555-48	Sequence 48, App
36	44	100.0	108	2	US-08-428-197-4	Sequence 4, App
37	44	100.0	108	5	PCT-US93-10555-4	Sequence 4, App
38	44	100.0	109	1	US-07-942-245-17	Sequence 17, App
39	44	100.0	109	1	US-07-942-245-21	Sequence 21, App
40	44	100.0	109	2	US-08-428-197-3	Sequence 3, App
41	44	100.0	109	5	PCT-US93-10555-3	Sequence 3, App
42	44	100.0	110	1	US-08-211-202-117	Sequence 117, App
43	44	100.0	110	3	US-08-545-809A-129	Sequence 129, App
44	44	100.0	111	1	US-08-211-202-134	Sequence 134, App
45	44	100.0	112	1	US-08-211-202-133	Sequence 133, App

ALIGNMENTS

RESULT 1
US-08-331-398A-66
Sequence 66, Application US/08331398A
Patent No. 5608039
GENERAL INFORMATION:
APPLICANT: Pastan, Ira
APPLICANT: Willingham, Mark
APPLICANT: Fitzgerald, David
APPLICANT: Brinkmann, Ulrich
APPLICANT: Pai, Lee
TITLE OF INVENTION: Single Chain B3 Antibody Fusion Proteins
TITLE OF INVENTION: and Their Uses (as amended)
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Stewart Street Plaza
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105-1492
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,398A
FILING DATE: 28-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/767,331
FILING DATE: 30-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/596,289
FILING DATE: 12-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 015280-126110US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-331-398A-66

Query Match 100.0%; Score 44; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.094;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVOPG 8
|||||||
DB 8 GGGVOPG 15

RESULT 2

US-08-652-558-42
Sequence 42, Application US/08652558
Patent No. 5861155
GENERAL INFORMATION:
APPLICANT: LIN, AUGUSTINE YEE-THARN
TITLE OF INVENTION: HUMANIZED ANTIBODIES AND USES
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER & WITCOFF
STREET: 75 STATE STREET, 23RD FLOOR
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,558
FILING DATE: JUNE 6, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB94/00387
FILING DATE: NOVEMBER 21, 1994
ATTORNEY/AGENT INFORMATION:
NAME: YANKWICH, LEON R.
REGISTRATION NUMBER: 30,237
REFERENCE/DOCKET NUMBER: 95,497-L
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-652-558-42

Query Match 100.0%; Score 44; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.094;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVOPG 8
|||||||
DB 8 GGGVOPG 15

RESULT 3
US-08-331-397B-66
Sequence 66, Application US/08331397B
Patent No. 5981726
GENERAL INFORMATION:
APPLICANT: Pastan, Ira
APPLICANT: Benhar, Itai
TITLE OF INVENTION: Chimeric and Mutationally Stabilized Tumor-Specific Antibody Fragments, Fusion Proteins, and Uses

TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Stewart Street Plaza
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105-1492

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,397B
FILING DATE: 28-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/767,331
FILING DATE: 30-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/596,289
FILING DATE: 12-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 015280-12612005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-331-397B-66

Query Match 100.0%; Score 44; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.094;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVOPG 8
|||||||
DB 8 GGGVOPG 15

RESULT 4
US-08-759-804A-65
Sequence 65, Application US/08759804A
Patent No. 5990296
GENERAL INFORMATION:
APPLICANT: Pastan, Ira
APPLICANT: Willingham, Mark
APPLICANT: Fitzgerald, David J.
APPLICANT: Brinkmann, Ulrich

TITLE OF INVENTION: Tumor-Specific Antibody Fragments,
Fusion Proteins, and Uses Thereof
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/759,804A
FILING DATE: 03-DEC-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/331,398
FILING DATE: 28-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/767,331
FILING DATE: 30-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/596,289
FILING DATE: 12-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen L.
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 015280-126140US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-759-804A-65

Query Match 100.0%; Score 44; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.094;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGGVQPG 8
| | | | | | | |
Db 8 GGGVQPG 15

RESULT 5
US-08-331-398A-55
Sequence 55, Application US/08331398A
Patent No. 5608039
GENERAL INFORMATION:
APPLICANT: Pastan, Ira
APPLICANT: Williamson, Mark
APPLICANT: Fitzgerald, David
APPLICANT: Brinkmann, Ulrich
APPLICANT: Pal, Lee
TITLE OF INVENTION: Single Chain B3 Antibody Fusion Proteins
TITLE OF INVENTION: and Their Uses (as amended)
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Stewart Street Plaza
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105-1492
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,398A
FILING DATE: 28-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/767,331
FILING DATE: 30-SEP-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/596,289
FILING DATE: 12-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 015280-126110US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-331-398A-55

Query Match 100.0%; Score 44; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGGVQPG 8
| | | | | | | |
Db 8 GGGVQPG 15

RESULT 6
US-08-401-908-14
Sequence 14, Application US/08401908
Patent No. 5684146
GENERAL INFORMATION:
APPLICANT: Yoshinobu OKUNO et al.
TITLE OF INVENTION: A DNA WHICH CODES FOR THE VARIABLE REGION OF
TITLE OF INVENTION: ANTIBODY
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/401,908
FILING DATE: March 10, 1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 18
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE: heavy chain of C179

US-08-401-908-14

Query Match 100.0%; Score 44; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQPG 8
|||||||
Db 8 GGGVQPG 15

RESULT 7

US-08-331-397B-55
; Sequence 55 Application US/0831397B
; Patent No. 5981726
; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira
; APPLICANT: Benhar, Itai
; TITLE OF INVENTION: Chimeric and Mutationally Stabilized Tumor-
; TITLE OF INVENTION: Specific Antibody Fragments, Fusion Proteins, and Uses
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Stewart Street Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/331.397B
; FILING DATE: 28-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/767,331
; FILING DATE: 30-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/596,289
; FILING DATE: 12-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 015280-126120US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-331-397B-55

Query Match 100.0%; Score 44; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQPG 8
|||||||
Db 8 GGGVQPG 15

RESULT 8
US-08-759-804A-54

; Sequence 54, Application US/08759804A

; Patent No. 5990296
; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira
; APPLICANT: Willingham, Mark
; APPLICANT: Fitzgerald, David J.
; APPLICANT: Brinkmann, Ulrich
; APPLICANT: Pal, Lee
; TITLE OF INVENTION: Tumor-Specific Antibody Fragments,
; TITLE OF INVENTION: Fusion Proteins, and Uses thereof
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/759.804A
; FILING DATE: 03-DEC-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/331,398
; FILING DATE: 28-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/767,331
; FILING DATE: 30-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/596,289
; FILING DATE: 12-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen L.
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 015280-126140US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-759-804A-54

Query Match 100.0%; Score 44; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQPG 8
|||||||
Db 8 GGGVQPG 15

RESULT 9

US-08-859-931A-2
; Sequence 2, Application US/08859931A
; Patent No. 5945510
; GENERAL INFORMATION:
; APPLICANT: PASANO, Alessio
; TITLE OF INVENTION: SUBSTANTIALLY PURE ZOUULIN, A
; TITLE OF INVENTION: PHYSIOLOGICAL MODULATOR OF
; TITLE OF INVENTION: MAMMALIAN TIGHT JUNCTIONS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:

```

ADDRESS: SUCHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W., Suite 800
CITY: Washington, D.C.
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,931A
FILING DATE: 21 MAY 1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: KIT, Gordon
REGISTRATION NUMBER: 30,764
REFERENCE/DOCKET NUMBER: A-6901
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Peptide
HYPOTHETICAL: NO
US-08-859-931A-2

```

```

Query Match      100.0%; Score 44; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGVQPG 8
   |||||
Db 8 GGGVQPG 15

```

```

RESULT 10
US-08-053-451B-113
; Sequence 113, Application US/08053451B
; Patent No. 5955584
; GENERAL INFORMATION:
; APPLICANT: Chen, Francis W.
; APPLICANT: Dittlow, Charles C.
; APPLICANT: Calenoff, Emanuel
; TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC
; NUMBER OF SEQUENCES: 176
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/053,451B
; FILING DATE: 26-APR-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 7606-033-999
; TELECOMMUNICATION INFORMATION:

```

```

TELEPHONE: 415-854-3660
TELEFAX: 415-854-3694
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2..3
OTHER INFORMATION: /note="Where Xaa = k or q = Lys or
; OTHER INFORMATION: Gln"
US-08-053-451B-113

```

```

Query Match      100.0%; Score 44; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGVQPG 8
   |||||
Db 7 GGGVQPG 14

```

```

RESULT 11
US-08-471-780C-80
; Sequence 80, Application US/08471780C
; Patent No. 5759808
; GENERAL INFORMATION:
; APPLICANT: Casterman, Cecile
; APPLICANT: Hamers, Raymond
; TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,780C
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/106,944
; FILING DATE: 17-AUG-1993
; APPLICATION NUMBER: FR 92402326.0
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 93401310.3
; FILING DATE: 21-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Potler, Jane E.R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 04958,0008-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Camelus dromedarius
US-08-471-780C-80

Query Match 100.0%; Score 44; DB 1; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQPG 8
DB 8 GGGVQPG 15

RESULT 12
US-08-467-282B-80
Sequence 80, Application US/08467282B
Patent No. 5800988
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,282B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Camelus dromedarius
US-08-467-282B-80

Query Match 100.0%; Score 44; DB 1; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQPG 8
DB 8 GGGVQPG 15

DB 8 GGGVQPG 15

RESULT 13
US-08-471-282A-80
Sequence 80, Application US/08471282A
Patent No. 5840853
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,282A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Camelus dromedarius
US-08-471-282A-80

Query Match 100.0%; Score 44; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQPG 8
DB 8 GGGVQPG 15

RESULT 14
US-08-466-710C-80
Sequence 80, Application US/08466710C
Patent No. 5874541
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,710C
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Camelus dromedarius
US-08-466-710C-80

Query Match          100.0%; Score 44; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 GGGLVQPG 8
        |||||||
DB      8 GGGLVQPG 15

RESULT 15
US-08-468-739C-80
; Sequence 80 Application US/08468739C
; Patent NO. 6015695
; GENERAL INFORMATION:
; APPLICANT: Castleman, Cecile
; APPLICANT: Hamers, Raymond
; TITLE OF INVENTION: Immunoglobulins devoid of light chains
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,739C
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Camelus dromedarius
US-08-468-739C-80

Query Match          100.0%; Score 44; DB 3; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 GGGLVQPG 8
        |||||||
DB      8 GGGLVQPG 15
```

Search completed: June 13, 2001, 14:27:06
Job time: 627 sec

Wed Jun 13 15:00:21 2001

pct-us01-05825a-23.ra1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: June 13, 2001, 14:25:44 ; Search time 150.28 Seconds
(without alignments)
3.043 Million cell updates/sec

Title: PCT-US01-05825A-24

Perfect score: 43

Sequence: 1 GGGLVODG 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.GeneSeq_0401.*
1: /SID56/gcgdata/geneSeq/geneSeq/AA1980.DAT.*
2: /SID56/gcgdata/geneSeq/geneSeq/AA1981.DAT.*
3: /SID56/gcgdata/geneSeq/geneSeq/AA1982.DAT.*
4: /SID56/gcgdata/geneSeq/geneSeq/AA1983.DAT.*
5: /SID56/gcgdata/geneSeq/geneSeq/AA1984.DAT.*
6: /SID56/gcgdata/geneSeq/geneSeq/AA1985.DAT.*
7: /SID56/gcgdata/geneSeq/geneSeq/AA1986.DAT.*
8: /SID56/gcgdata/geneSeq/geneSeq/AA1987.DAT.*
9: /SID56/gcgdata/geneSeq/geneSeq/AA1988.DAT.*
10: /SID56/gcgdata/geneSeq/geneSeq/AA1989.DAT.*
11: /SID56/gcgdata/geneSeq/geneSeq/AA1990.DAT.*
12: /SID56/gcgdata/geneSeq/geneSeq/AA1991.DAT.*
13: /SID56/gcgdata/geneSeq/geneSeq/AA1992.DAT.*
14: /SID56/gcgdata/geneSeq/geneSeq/AA1993.DAT.*
15: /SID56/gcgdata/geneSeq/geneSeq/AA1994.DAT.*
16: /SID56/gcgdata/geneSeq/geneSeq/AA1995.DAT.*
17: /SID56/gcgdata/geneSeq/geneSeq/AA1996.DAT.*
18: /SID56/gcgdata/geneSeq/geneSeq/AA1997.DAT.*
19: /SID56/gcgdata/geneSeq/geneSeq/AA1998.DAT.*
20: /SID56/gcgdata/geneSeq/geneSeq/AA1999.DAT.*
21: /SID56/gcgdata/geneSeq/geneSeq/AA2000.DAT.*
22: /SID56/gcgdata/geneSeq/geneSeq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	43	100.0	8	21	Y79128
2	38	88.4	8	21	Y79126
3	37	86.0	119	17	R97334
4	37	86.0	334	20	Y29904
5	37	86.0	338	22	B65728
6	37	86.0	339	22	Y29905
7	37	86.0	348	20	Y29903
8	36	83.7	8	21	Y79127
9	36	83.7	13	20	W86096
10	36	83.7	13	20	W86076
11	36	83.7	13	20	W86088

12	36	83.7	18	16	W73953	IGG heavy chain fr
13	36	83.7	18	16	R82833	N-terminal of C179
14	36	83.7	20	20	W94487	Human adult heart
15	36	83.7	20	21	Y79130	Human adult heart
16	36	83.7	30	17	R87049	Human group III he
17	36	83.7	30	21	Y68811	Human heavy chain
18	36	83.7	40	7	P61028	H-chain variable r
19	36	83.7	62	16	R76969	HSV-neuraltising a
20	36	83.7	73	21	Y64774	Human 5' EST relat
21	36	83.7	84	21	Y64925	Human 5' EST relat
22	36	83.7	87	17	R92997	Homologous sequenc
23	36	83.7	93	20	Y11665	Human 5' EST secre
24	36	83.7	96	21	Y56646	Partial peptide fr
25	36	83.7	96	21	Y64844	Human 5' EST relat
26	36	83.7	97	21	B40089	Anti-H1L12 antibod
27	36	83.7	97	21	B40092	Anti-H1L12 antibod
28	36	83.7	97	21	B40112	Anti-H1L12 antibod
29	36	83.7	97	21	B40135	Anti-H1L12 antibod
30	36	83.7	98	14	R34279	Human TNF binding
31	36	83.7	98	15	R52066	Heavy chain variab
32	36	83.7	98	15	R54816	SpA-reactive VH re
33	36	83.7	98	16	R72074	DP54 VH region. H
34	36	83.7	98	19	W59614	Anti-RSV F protein
35	36	83.7	98	21	B40072	Anti-H1L12 antibod
36	36	83.7	98	21	B40084	Anti-H1L12 antibod
37	36	83.7	98	21	B40085	Anti-H1L12 antibod
38	36	83.7	98	21	B40086	Anti-H1L12 antibod
39	36	83.7	98	21	B40087	Anti-H1L12 antibod
40	36	83.7	98	21	B40093	Anti-H1L12 antibod
41	36	83.7	98	21	B40094	Anti-H1L12 antibod
42	36	83.7	98	21	B40095	Anti-H1L12 antibod
43	36	83.7	98	21	B40096	Anti-H1L12 antibod
44	36	83.7	98	21	B40097	Anti-H1L12 antibod
45	36	83.7	98	21	Y56660	Partial peptide fr

ALIGNMENTS

RESULT 1	
ID Y79128	standard; Peptide: 8 AA.
XX	
AC Y79128;	
XX	
DT 05-JUN-2000	(first entry)
XX	
DE	Peptide antagonist of zonulin.
XX	
KW	zonulin; antagonist; zonula occludens toxin receptor;
KW	blood-brain barrier; antiinflammatory; cerebroprotective;
KW	neuroprotective; dermatological; antitumor; antiviral;
KW	antibacterial; cytostatic; anti-HIV; vulnereary; antiallergic;
KW	hypotensive; immunosuppressive; antiparasitic; vasotropic;
KW	gastrointestinal inflammation; therapy.
OS	Synthetic.
XX	
PN	WO200007609-A1.
XX	
PD	17-FEB-2000.
XX	
PF	28-JUL-1999; 99WO-US16683.
XX	
PR	03-AUG-1998; 98US-0127815.
XX	
PA	(UYMA-) UNIV MARYLAND BALTIMORE.
XX	
PI	Fasano A;
XX	
DR	WPI; 2000-205565/18.
XX	
PT	New peptide antagonist of zonulin useful as antiinflammatory agent for

PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 XX shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 PS Claim 1; Page 48; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infection, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC celiac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.

Sequence 8 AA;

Query Match 100.0%; Score 43; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGIVDQG 8
 DB 1 gggivqdq 8

RESULT 2
 Y79126
 XX Y79126 standard; Peptide; 8 AA.
 AC Y79126;

DT 05-JUN-2000 (first entry)

DE Peptide antagonist of zonulin.

KW Zonulin: antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; antiviral;
 KW antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.

OS Synthetic.

WO200007609-A1.

PD 17-FEB-2000.

PF 28-JUL-1999; 99WO-US16683.

PR 03-AUG-1998; 98US-0127815.

PA (UYMA-) UNIV MARYLAND BALTIMORE.

PI Fasano A;
 XX WI; 2000-205565/18.

PT New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 PS Claim 1; Page 48; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infection, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC celiac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.

Sequence 8 AA;

Query Match 88.4%; Score 38; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGIVDQG 8
 DB 1 gggivqdq 8

RESULT 3
 R97334
 ID R97334 standard; Protein; 119 AA.
 AC R97334;

DT 15-OCT-1996 (first entry)

DE Murine NEMM MN-14 variable region heavy chain.

KW Monoclonal antibody; humanised; mouse; framework region; FR; CDR;
 KW complementary determining region; anti-carcinoembryonic antigen;
 KW CEA; diagnosis; imaging; therapy; immune response.

OS Mus musculus.

Key Location/Qualifiers

FT Binding-site 31..35

FT Binding-site /label= CDR 1.

FT Binding-site 50..66

FT Binding-site /label= CDR 2.

FT Binding-site 99..108

FT Binding-site /label= CDR 3.

PM WO9611013-A1.
XX
PD 18-APR-1996.
XX
XX 28-SEP-1995; 95WO-US11964.
XX
XX 05-OCT-1994; 94US-0318157.
XX
PA (IMMU-) IMMUNOMEDICS INC.
XX
XX Armour KL, Hansen HJ;
XX
DR WPI: 1996-209653/21.
DR N-PSDB; T29010.
XX
XX New humanised anti-CEA monoclonal antibody - having engrafted murine
PT CDRs, used for diagnosis, imaging and therapy of CEA-producing
PT cancers
XX
XX Example 5; Figure 1; 62pp; English.
XX
XX New humanised monoclonal antibodies (Mabs) comprising the
CC complementary determining regions (CDRs) of a parental murine class
CC III anti-carcinembryonic (CEA) MAb engrafted to the framework
CC regions (FRs) of a heterologous antibody which can be derived from
CC any species including human, retain the anti-CEA binding specificity
CC of the parental murine MAb but are less immunogenic in a human
CC subject than the parental MAb. The humanised antibodies can be used
CC in diagnosis, imaging and therapy of CEA-producing cancers and
CC patients receiving the humanised antibodies and conjugates show
CC improved therapeutic results, decreased immune responses and
CC decreased immune-mediated adverse effects compared to the parent
CC antibody. See R87313-97333 for the CDR sequence and framework
CC regions of the humanised Mabs.
XX
XX Sequence 119 AA;
SQ

Query Match 86.0%; Score 37; DB 17; Length 119;
Best Local Similarity 87.5%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
| | | | | | |
DB 8 g9g1vqsg 15

RESULT 4
Y29904
ID Y29904 standard; Protein: 334 AA.
XX
AC Y29904;
XX
DT 17-NOV-1999 (first entry)
XX
DE Human MDC and human scFv fusion protein.
XX
XX Chemokine; tumour; viral; antigen; fusion protein; cancer; vaccine;
KW Immune response; HIV; Infection.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX WO9946392-A1.
PN
XX
PD 16-SEP-1999.
XX
PF 12-MAR-1999; 99WO-US05345.
XX
PR 12-MAR-1998; 98US-0077745.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX

PI Kwak LW, Biragyn A;
XX
XX WPI: 1999-551418/46.
DR
XX
XX New fusion polypeptides comprising a chemokine and a tumour antigen or
PT HIV antigen, used for treating cancers or treating or preventing HIV
PT infection
XX
XX Claim 73; Page 134-135; 142pp; English.
PS
XX
XX The present invention describes fusion proteins comprising a chemokine
CC and a tumour antigen or HIV antigen. Specifically claimed fusion proteins
CC comprise: (1) human monocyte chemoattractant protein-3 (MCP-3) and human
CC Muc-1; (2) human interferon-induced protein 10 (IP-10) and human Muc-1;
CC (3) human macrophage-derived chemokine (MDC) and human Muc-1; (4) human
CC SDF-1 and human Muc-1; (5) human IP-10 and HIV gp120; (6) human MCP-3 and
CC HIV gp120; (7) human MDC and HIV gp120; and (8) human SDF-1 and HIV
CC gp120. The fusion proteins, and nucleotide sequences encoding them, can
CC be used for producing an immune response, e.g. an effector T cell immune
CC response. They can also be used for treating cancer or treating or
CC preventing HIV infection. The fusion proteins and/or nucleotide sequences
CC can be used in in vitro diagnostic assays, as well as in screening assays
CC for identifying unknown tumour antigen epitopes and fine mapping of
CC tumour antigen epitopes. The present sequence represents a specifically
CC claimed fusion protein from the present invention.
XX
XX Sequence 334 AA;
SQ

Query Match 86.0%; Score 37; DB 20; Length 334;
Best Local Similarity 87.5%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
| | | | | | |
DB 87 g9g1vqsg 94

RESULT 5
B65728
ID B65728 standard; Protein: 338 AA.
XX
AC B65728;
XX
DT 27-MAR-2001 (first entry)
XX
DE Nucellin-like aspartic protease #1.
XX
XX Cell death modulator; programmed cell death; PCD; apoptosis;
KW forestry plant.
XX
XX Pinus radiata.
OS
XX
PN WO200075331-A1.
XX
PD 14-DEC-2000.
XX
XX 02-JUN-2000; 2000WO-N200086.
PF
XX
PR 04-JUN-1999; 99US-0325932.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
XX Flinn B, Lasham A;
PI
XX
DR WPI: 2001-061724/07.
DR N-PSDB; F44734.
XX
XX Novel defender against cell death polynucleotide useful for modulating
PT programmed cell death pathway and specific development pathways in
PT forestry plant -
XX

PS Claim 22; Page 69; 142pp; English.
 CC The present invention relates to coding sequences (see F44740-F44840 and
 CC F44843-F44844) and proteins (see B65714-B65814) involved in programmed
 CC cell death (PCD; apoptosis). The coding sequences and proteins of the
 CC present invention are useful for modulating a PCD or cell death pathway
 CC and various developmental pathways in a forestry plant, by stably
 CC incorporating one of the present coding sequences into the genome of the
 CC forestry plant, where the coding sequence provides a PCD pathway that is
 CC not present in a native form of the forestry plant.
 XX
 SQ Sequence 338 AA;

Query Match 86.0%; Score 37; DB 22; Length 338;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGUQD 7
 Db 285 9991vqgd 291

RESULT 6
 Y29905
 ID Y29905 standard; Protein: 339 AA.
 AC Y29905;
 XX
 DT 17-NOV-1999 (first entry)
 DE Human SDF-1 and human scfV fusion protein.
 KW Chemokine; tumour; viral; antigen; fusion protein; cancer; vaccine;
 KM Immune response; HIV; Infection.
 OS Homo sapiens.
 OS Synthetic.
 PN MO9946392-A1.
 PD 16-SEP-1999.
 PF 12-MAR-1999; 99WO-US05345.
 PR 12-MAR-1998; 98US-0077745.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Kwak LW, Biragyn A;
 XX WPI; 1999-551418/46.
 DR
 PT New fusion polypeptides comprising a chemokine and a tumour antigen or
 PT HIV antigen, used for treating cancers or treating or preventing HIV
 PT infection.
 XX
 PS Claim 73; Page 136; 142pp; English.
 XX
 CC The present invention describes fusion proteins comprising a chemokine
 CC and a tumour antigen or HIV antigen. Specifically claimed fusion proteins
 CC comprise: (1) human monocyte chemoattractant protein-3 (MCP-3) and human
 CC Muc-1; (2) human interferon-induced protein 10 (IP-10) and human Muc-1;
 CC (3) human macrophage-derived chemokine (MDC) and human Muc-1; (4) human
 CC SDF-1 and human Muc-1; (5) human IP-10 and HIV gp120; (6) human MCP-3 and
 CC HIV gp120; (7) human MDC and HIV gp120; and (8) human SDF-1 and HIV
 CC gp120. The fusion proteins, and nucleotide sequences encoding them, can
 CC be used for producing an immune response, e.g. an effector T cell immune
 CC response. They can also be used for treating cancer or treating or
 CC preventing HIV infection. The fusion proteins and/or nucleotide sequences
 CC can be used in vitro diagnostic assays, as well as in screening assays
 CC for identifying unknown tumour antigen epitopes and fine mapping of
 CC tumour antigen epitopes. The present sequence represents a specifically

CC claimed fusion protein from the present invention.
 XX
 SQ Sequence 339 AA;

Query Match 86.0%; Score 37; DB 20; Length 339;
 Best Local Similarity 87.5%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGUQD 8
 Db 92 9991vqsg 99

RESULT 7
 Y29903
 ID Y29903 standard; Protein: 348 AA.
 AC Y29903;
 XX
 DT 17-NOV-1999 (first entry)
 DE Human MCP-3 and human scfV fusion protein.
 KW Chemokine; tumour; viral; antigen; fusion protein; cancer; vaccine;
 KM Immune response; HIV; Infection.
 OS Homo sapiens.
 OS Synthetic.
 PN MO9946392-A1.
 PD 16-SEP-1999.
 PF 12-MAR-1999; 99WO-US05345.
 PR 12-MAR-1998; 98US-0077745.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Kwak LW, Biragyn A;
 XX WPI; 1999-551418/46.
 DR
 PT New fusion polypeptides comprising a chemokine and a tumour antigen or
 PT HIV antigen, used for treating cancers or treating or preventing HIV
 PT infection.
 XX
 PS Claim 73; Page 132; 142pp; English.
 XX
 CC The present invention describes fusion proteins comprising a chemokine
 CC and a tumour antigen or HIV antigen. Specifically claimed fusion proteins
 CC comprise: (1) human monocyte chemoattractant protein-3 (MCP-3) and human
 CC Muc-1; (2) human interferon-induced protein 10 (IP-10) and human Muc-1;
 CC (3) human macrophage-derived chemokine (MDC) and human Muc-1; (4) human
 CC SDF-1 and human Muc-1; (5) human IP-10 and HIV gp120; (6) human MCP-3 and
 CC HIV gp120; (7) human MDC and HIV gp120; and (8) human SDF-1 and HIV
 CC gp120. The fusion proteins, and nucleotide sequences encoding them, can
 CC be used for producing an immune response, e.g. an effector T cell immune
 CC response. They can also be used for treating cancer or treating or
 CC preventing HIV infection. The fusion proteins and/or nucleotide sequences
 CC can be used in vitro diagnostic assays, as well as in screening assays
 CC for identifying unknown tumour antigen epitopes and fine mapping of
 CC tumour antigen epitopes. The present sequence represents a specifically
 CC claimed fusion protein from the present invention.
 XX
 SQ Sequence 348 AA;

Query Match 86.0%; Score 37; DB 20; Length 348;
 Best Local Similarity 87.5%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGIVODG 8
 DB 101 gggivqpg 108

RESULT 8
 ID Y79127 standard; Peptide: 8 AA.
 Y79127:
 AC Y79127:
 XX
 XX
 DT 05-JUN-2000 (first entry)
 XX
 DE Peptide antagonist of zonulin.

XX Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; antilucer; antiviral;
 KW antibacterial; cytostatic; anti-HIV; vulnerey; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.

OS Synthetic.

XX WO200007609-A1.

PD 17-FEB-2000.

PF 28-JUL-1999; 99WO-US16683.

PR 03-AUG-1998; 98US-0127815.

PA (UYMA-) UNITV MARYLAND BALTIMORE.

PI Faasano A;

XX MPI; 2000-205565/18.

XX New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -

PS Claim 1: Page 48; 69pp; English.

XX This present sequence is that of a peptide antagonist of zonulin
 CC (Z), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.

XX Sequence 8 AA:

Query Match 83.7%; Score 36; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGIVODG 8
 DB 1 gggivqpg 8

RESULT 9
 ID W86096 standard; peptide: 13 AA.
 W86096:
 AC W86096:
 XX
 XX
 DT 03-MAR-1999 (first entry)
 XX
 DE Peptide from humanised A33 heavy chain variable region.

XX Non-immunogenic; epitope; T-cell; immunogenicity; immune system;
 KW immunoglobulin; therapeutic; streptokinase; diagnosis; human; A33.

XX Homo sapiens.

OS Homo sapiens.

XX WO9852976-A1.

PD 26-NOV-1998.

PF 21-MAY-1998; 98WO-GB01473.

PR 14-APR-1998; 98GB-0007751.

PR 21-MAY-1997; 97GB-0010480.

PR 31-JUL-1997; 97GB-0016197.

PR 28-NOV-1997; 97GB-0025270.

PR 02-DEC-1997; 97US-0067235.

PA (BIOV-) BIOVATION LTD.

PI Carr FJ;

XX MPI; 1999-045301/04.

XX Reducing immunogenicity of proteins - by modifying the amino acid
 PT sequence of the protein to eliminate potential epitopes for T-cells
 PT of a given species
 XX Example 5; Page 34; 77pp; English.

XX The invention relates to a method for the production of non-immunogenic
 CC proteins. The method comprises determining at least part of the amino
 CC acid sequence of the protein; (b) identifying in the amino acid sequence
 CC one or more potential epitopes for T-cells (T-cell epitopes) of the given
 CC species; and (c) modifying the amino acid sequence to eliminate at least
 CC one of the T-cell epitopes identified in step (b) thereby to eliminate or
 CC reduce the immunogenicity of the protein when exposed to the immune
 CC system of the given species. A method of analysing a pre-existing protein
 CC to predict the basis for immunogenic responses is also provided. The
 CC methods can be used particularly for reducing the immunogenicity of
 CC immunoglobulins or therapeutic proteins, e.g. streptokinase. The products
 CC can be used for diagnosis and therapy. Sequences W86075-W86114 represent
 CC peptides from humanised A33 heavy chain variable region predicted by
 CC peptide threading.

XX Sequence 13 AA:

Query Match 83.7%; Score 36; DB 20; Length 13;
 Best Local Similarity 87.5%; Pred. No. 2;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGIVODG 8
 DB 1 gggivqpg 8

Db 6 999lvpg 13

RESULT 10

ID W86076 standard; peptide; 13 AA.

AC W86076;

DT 03-MAR-1999 (first entry)

DE Peptide from humanised A33 heavy chain variable region.

KW Non-immunogenic; epitope; T-cell; immunogenicity; immune system; immunoglobulin; therapeutic; streptokinase; diagnosis; human; A33.

OS Homo sapiens.

PN W09852976-A1.

PD 26-NOV-1998.

PF 21-MAY-1998; 98WO-GB01473.

PR 14-APR-1998; 98GB-0007751.

PR 21-MAY-1997; 97GB-0010480.

PR 31-JUL-1997; 97GB-0016197.

PR 28-NOV-1997; 97GB-0025270.

PR 02-DEC-1997; 97US-0067235.

PA (BIOV-) BIOVATION LTD.

PI Carr FJ;

PT Reducing immunogenicity of proteins - by modifying the amino acid sequence of the protein to eliminate potential epitopes for T-cells

PS Example 5; Page 33; 77pp; English.

The invention relates to a method for the production of non-immunogenic proteins. The method comprises determining at least part of the amino acid sequence of the protein; (b) identifying in the amino acid sequence one or more potential epitopes for T-cells (T-cell epitopes) of the given species; and (c) modifying the amino acid sequence to eliminate at least one of the T-cell epitopes identified in step (b) thereby to eliminate or reduce the immunogenicity of the protein when exposed to the immune system of the given species. A method of analysing a pre-existing protein to predict the basis for immunogenic responses is also provided. The methods can be used particularly for reducing the immunogenicity of immunoglobulins or therapeutic proteins, e.g. Streptokinase. The products can be used for diagnosis and therapy. Sequences W86075-W86114 represent peptides from humanised A33 heavy chain variable region predicted by peptide threading.

SQ Sequence 13 AA;

Query Match 83.7%; Score 36; DB 20; Length 13;

Best Local Similarity 87.5%; Pred. NO. 2;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGIVODG 8

Db 6 999lvpg 13

RESULT 11

ID W86088

XX W86088 standard; peptide; 13 AA.

AC W86088;

DT 03-MAR-1999 (first entry)

DE Peptide from humanised A33 heavy chain variable region.

KW Non-immunogenic; epitope; T-cell; immunogenicity; immune system; immunoglobulin; therapeutic; streptokinase; diagnosis; human; A33.

OS Homo sapiens.

PN W09852976-A1.

PD 26-NOV-1998.

PF 21-MAY-1998; 98WO-GB01473.

PR 14-APR-1998; 98GB-0007751.

PR 21-MAY-1997; 97GB-0010480.

PR 31-JUL-1997; 97GB-0016197.

PR 28-NOV-1997; 97GB-0025270.

PR 02-DEC-1997; 97US-0067235.

PA (BIOV-) BIOVATION LTD.

PI Carr FJ;

PT Reducing immunogenicity of proteins - by modifying the amino acid sequence of the protein to eliminate potential epitopes for T-cells

PS Example 5; Page 33; 77pp; English.

The invention relates to a method for the production of non-immunogenic proteins. The method comprises determining at least part of the amino acid sequence of the protein; (b) identifying in the amino acid sequence one or more potential epitopes for T-cells (T-cell epitopes) of the given species; and (c) modifying the amino acid sequence to eliminate at least one of the T-cell epitopes identified in step (b) thereby to eliminate or reduce the immunogenicity of the protein when exposed to the immune system of the given species. A method of analysing a pre-existing protein to predict the basis for immunogenic responses is also provided. The methods can be used particularly for reducing the immunogenicity of immunoglobulins or therapeutic proteins, e.g. Streptokinase. The products can be used for diagnosis and therapy. Sequences W86075-W86114 represent peptides from humanised A33 heavy chain variable region predicted by peptide threading.

SQ Sequence 13 AA;

Query Match 83.7%; Score 36; DB 20; Length 13;

Best Local Similarity 87.5%; Pred. NO. 2;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGIVODG 8

Db 3 999lvpg 10

RESULT 12

ID W73953

XX W73953 standard; peptide; 18 AA.

AC W73953;

DT 26-APR-1999 (first entry)

DE Igg heavy chain fragment, Igg (VH) (18).

KW Igg heavy chain; immune complex removal; conglutinin; infection;

KM serum sickness; chronic immune complex nephritis; parasitic disease;
 KM autoimmune disease; neoplastic disease; immune pathogenesis; therapy.
 OS Synthetic.
 XX
 PN WO9506254-A1.
 XX
 PD 02-MAR-1995.
 XX
 PF 24-AUG-1994; 94WO-US09407.
 XX
 PR 24-AUG-1993; 93US-0111071.
 XX
 PA (IMMU-) APPLIED IMMUNE SCI INC.
 XX
 PI Lee YM, Odonoghue G, Okarma TB;
 XX
 DR WPI; 1995-106945/14.
 XX
 PT New biomedical device for binding immune complexes - comprising
 PT conglutinin covalently bound to a solid phase material
 PS
 XX Example 4; Fig 11; 58pp; English.
 CC This sequence represents a fragment of the IgG heavy
 CC chain. This sequence was isolated using the biomedical device of the
 CC invention, which comprises: (a) a biocompatible support; and
 CC (b) conglutinin which is covalently bound to the support. The device can
 CC be used for removing immune complexes (IC) from a fluid. This can be used
 CC for treating a mammal or for evaluating the status of a mammal. It can be
 CC used for treating diseases such as serum sickness, chronic immune complex
 CC nephritis, bacterial infections, viral infections, parasitic diseases,
 CC autoimmune diseases or neoplastic diseases. The devices can also be used
 CC for purification of IC and to study immune pathogenesis and antigens
 CC associated with a disease state. The covalently immobilised conglutinin
 CC can capture IC with high capacity and specificity and allows elution
 CC under mild conditions. The devices are stable with no loss of IC binding
 CC activity after repeated use.
 CC
 SQ Sequence 18 AA;
 CC
 Query Match 83.7%; Score 36; DB 16; Length 18;
 Best Local Similarity 87.5%; Pred. No. 2.8;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 GGGLVODG 8
 Db 8 99gltvpg 15
 DE N-terminal of C179 heavy chain.
 XX
 DE 09-MAY-1996 (first entry)
 DT
 XX
 AC R82833;
 XX
 DE 09-MAY-1996 (first entry)
 DT
 XX
 OS Mus musculus.
 XX
 PN EP675199-A2.
 XX
 PD 04-OCT-1995.
 XX
 PF 14-MAR-1995; 95EP-0301664.
 XX

PR 30-MAR-1994; 94JP-0082693.
 XX
 PA (TAKI) TAKARA SHUZO CO LTD.
 XX
 PI Atsushi O, Hirofumi Y, Ikunoshin K, Takashi T, Yoshinobu O;
 XX
 DR WPI; 1995-338286/44.
 XX
 PT Gene encoding variable region of anti-human influenza A type virus
 PT antibody - useful for prodn. of artificial antibodies
 PS
 XX Example 2; Page 24; 42pp; English.
 CC R82833 and R82834 represent the N-terminal sequences of the heavy chain
 CC and light chain of an anti-human influenza A type virus antibody. The
 CC antibody these sequences were obtained from was isolated from hydropoma
 CC C179. The antibody recognises the stem region of the haemagglutinin (HA)
 CC molecule of the H1N1 and H2N2 subtypes of human influenza A type virus,
 CC and shows neutralisation activity against these two subtypes. The
 CC antibody shows no recognition of the H2N2 subtype. The DNA encoding the
 CC variable heavy chain sequence (see T04159) and the variable light chain
 CC coding sequence (see T04160) are useful in the production of artificial
 CC antibodies and antigen-binding polypeptides. These artificial antibodies
 CC and polypeptides are useful in the diagnosis and treatment of human
 CC influenza. As the antibodies recognise the stem region of the HA
 CC molecule, the influenza virus will be recognised even if the HA molecule
 CC changes. This provides an advantage over current vaccines, as the virus
 CC periodically alters it's HA molecule.
 SQ Sequence 18 AA;
 CC
 Query Match 83.7%; Score 36; DB 16; Length 18;
 Best Local Similarity 87.5%; Pred. No. 2.8;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 GGGLVODG 8
 Db 8 99gltvpg 15

RESULT 14
 W94487
 ID W94487 standard; peptide: 20 AA.
 XX
 AC W94487;
 XX
 DE 21-APR-1999 (first entry)
 DT
 XX
 DE Human adult heart zonulin N-terminal peptide.
 XX
 KM Zonulin; mammalian tight junction; zonula occludens toxin; ZOT;
 KM Vibrio cholerae; vaccine; cholera toxin; polyclonal antibody;
 KM intestinal mucosa; nasal mucosa; blood brain barrier.
 XX
 OS Homo sapiens.
 XX
 PN WO9852415-A1.
 XX
 PD 26-NOV-1998.
 XX
 PF 28-APR-1998; 98WO-US07636.
 XX
 PR 21-MAY-1997; 97US-0859931.
 XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Fasano A;
 XX
 DR WPI; 1999-070123/06.
 XX
 PT New purified zonulin - which is capable of reversibly opening
 PT mammalian tight junctions, used for enhancing the delivery of agents

The present invention describes pure zonulin which has an apparent molecular weight of 47 kD, as determined by SDS-PAGE, which is recognised by both anti-tan polyclonal antibody and by anti-zonula occludens toxin (ZOT) polyclonal antibody, and is capable of reversibly opening mammalian tight junctions. Zonulin proteins function as physiological modulators of mammalian tight junctions. They can be used for enhancing the absorption of therapeutic agents across tight junctions of intestinal and nasal mucosa and across tight junctions of the blood brain barrier. Zonulin can be used with agents such as drugs, e.g. lidocaine, adenosine, dibucamine, dopamine, epinephrine, norepinephrine, phenoltolamine, dexampram, alfentanil, deszoin, nalbuphine, buprenorphine, naloxone, ketorolac, midazolam, propofol, metacurine, vincarium, succinylcholine, cycarblade, miltownin doxorublin, vincristine, vinblastine, methicillin, mezlocillin, piperacillin, cefotaxim, ceftriaxone, cefmetazole and aztreonam, a hormone e.g. testosterone, nandrolone, memotropins, insulin, urofollitropin, interferon-alpha, interferon-beta, interferon-gamma, interleukin-1 (IL-1), IL-2, IL-4, IL-8, polyvalent IgG, specific IgG, IgA, or IgM. The proteins can also be used for the production of antibodies which can be used to assay for zonulin in body tissue or fluids, or in affinity-purification of zonulin. The present sequence represents an N-terminal peptide of zonulin.

Query Match	83.7%	Score 36;	DB 20;	Length 20;
Best Local Similarity	87.5%;	Pred. No. 3.2;		
Matches	7;	Conservative	0;	Mismatches 1; Indels 0; Gaps 0

QY . 1 GGGLVQDG 8 .
| | | | | | |
pb 8 gggLVqpg 15

Y79130 standard; Peptide; 20 AA.

05-JUN-2000 (first entry)

human; blood-brain barrier; anti-inflammation; zoonosis; antagonist; zoonosis; occludens toxin receptor;

gastrointestinal inflammation; therapy.

Homo sapiens.

WO2000007609-A1

17-FEB-2000.

28-JUL-1999; 99WO-US16683.

03-AUG-1998; 98US-0127815.

(OYMA-) UNIV MARYLAND BALTIMORE.

Fasano A;

WPI; 2000-205565/18.

new peptide antagonist of zonulin useful as antiinflammatory agent for treating cerebral ischemia, stroke, cerebral edema, gastritis, shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -

The present sequence is that of the N-terminal region of adult human heart zonulin. The N-terminal sequences of human adult and zoonal zonulins (see W79130-35) were compared with Vibrio cholerae zonula occludens toxin (ZOT) to identify a common motif thought to be involved in receptor binding. Peptide antagonists (see W79105-29) based on this motif are useful as antiinflammatory agents for treatment of gastrointestinal inflammation, and for treatment of conditions associated with breakdown of the blood-brain barrier.

Query Match	83.7%	Score 36;	DB 21;	Length 20;
Best Local Similarity	87.5%	Pred. No. 3.2;		
Matches	7;	Conservative	0;	Mismatches 1; Indels 0; Gaps 0;

QY	1	GGGLVQDG	8
Db	8	ggglvqpg	15

Search completed: June 13, 2001, 14:25:45
Job time: 664 sec

Job time: 664 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:23:10 ; Search time 87.97 seconds
(without alignments)
6.250 Million cell updates/sec

Title: PCT-US01-05825A-24

Perfect score: 43

Sequence: 1 GCGLVQDG 8

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR67:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	88.4	1000	2 C82630	serine proteinase
2	37	86.0	101	2 PU0001	Ig heavy chain V r
3	37	86.0	123	2 C30556	Ig heavy chain V r
4	37	86.0	123	2 D30539	Ig heavy chain V r
5	37	86.0	146	2 I47184	Ig heavy chain var
6	36	83.7	36	2 D31485	Ig heavy chain V r
7	36	83.7	38	2 S33402	Ig heavy chain V r
8	36	83.7	59	2 A27606	Ig heavy chain V-a
9	36	83.7	59	2 S36381	Ig heavy chain V r
10	36	83.7	60	2 S36382	Ig heavy chain V r
11	36	83.7	82	2 C34964	Ig heavy chain pre
12	36	83.7	94	2 D25913	Ig heavy chain V r
13	36	83.7	94	2 I67528	CD33 antigen homol
14	36	83.7	95	2 I67527	CD33 antigen homol
15	36	83.7	97	2 PH0872	Ig heavy chain V r
16	36	83.7	97	2 S26885	Ig heavy chain V r
17	36	83.7	97	2 S26886	Ig heavy chain V r
18	36	83.7	97	2 S26890	Ig heavy chain V r
19	36	83.7	97	2 S46462	Ig heavy chain V r
20	36	83.7	97	2 S26895	Ig heavy chain V r
21	36	83.7	97	2 S54855	Ig heavy chain V r
22	36	83.7	98	2 PL0121	Ig heavy chain V-I
23	36	83.7	98	2 PH0874	Ig heavy chain V r
24	36	83.7	98	2 PL0123	Ig heavy chain V-I
25	36	83.7	98	2 S26896	Ig heavy chain V r
26	36	83.7	98	2 S26945	Ig heavy chain V r
27	36	83.7	98	2 S26927	Ig heavy chain V r
28	36	83.7	98	2 S26932	Ig heavy chain V r
29	36	83.7	98	2 S26891	Ig heavy chain V r

30	36	83.7	98	2 S26894	Ig heavy chain V r
31	36	83.7	98	2 S26889	Ig heavy chain V r
32	36	83.7	98	2 S26933	Ig heavy chain V r
33	36	83.7	98	2 S26934	Ig heavy chain V r
34	36	83.7	100	2 PL0122	Ig heavy chain V-I
35	36	83.7	100	2 S69896	Ig heavy chain V r
36	36	83.7	100	2 S26925	Ig heavy chain V r
37	36	83.7	100	2 S26926	Ig heavy chain V r
38	36	83.7	101	2 PU0003	Ig heavy chain V r
39	36	83.7	101	2 PU0002	Ig heavy chain V r
40	36	83.7	105	2 S38488	Ig heavy chain - h
41	36	83.7	105	2 PL0255	Ig heavy chain V r
42	36	83.7	106	2 S24521	Ig heavy chain V r
43	36	83.7	111	1 MHM576	Ig heavy chain V-I
44	36	83.7	111	2 PH1028	Ig heavy chain V r
45	36	83.7	111	2 S69911	Ig V-D-J region (R

ALIGNMENTS

RESULT 1
C82630 serine proteinase XF1851 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: C82630
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; PMID:20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: C82630
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1000 <SIM>
A:Cross-references: GB:AE004006; GB:AE003849; NID:g9106932; PIDN:AAF84657.1; GSPDB:GN
A:Experimental source: strain 9a5c
R:Stimpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitchima, J.P.; Krieger, J.E.; Kurano, E.F.; La
chido, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sava
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1851

Query Match 88.4%; Score 38; DB 2; Length 1000;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCGLVQDG 8
Db 464 GCGLVQDG 471

RESULT 2
PU0001 Ig heavy chain V region (V1) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 06-Jun-1997
C:Accession: PU0001
R:Ferguson, S.E.; Rudikoff, S.; Osborne, B.A.
J. Exp. Med. 168, 1339-1349, 1988

A:Title: Interaction and sequence diversity among T15 VH genes in CBA/J mice.
A:Reference number: PU0001; MUID:89010532
A:Accession: PU0001
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-101 <PER>
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:15-100/Domain: Immunoglobulin homology <IMM>

Query Match
Best Local Similarity 86.0%; Score 37; DB 2; Length 101;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
|||||
Db 8 GGGLVQSG 15

RESULT 3

Ig heavy chain V regions (7C9, 6D10, and 2G6) - mouse
C:Species: Mus musculus (house mouse)

C:Date: 23-Feb-1989 #sequence_revision 23-Feb-1989 #text_change 16-Aug-1996
C:Accession: C30556

R:Clatlin, J.L.; Berry, J.; Flaherty, D.; Dunnick, W.

J. Immunol. 138, 3060-3068, 1987

A:Title: Somatic evolution of diversity among anti-phosphocholine antibodies induced with A:Reference number: A30556; MUID:87196439

A:Accession: C30556

A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-123 <CLA>

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:15-100/Domain: Immunoglobulin homology <IMM>

Query Match
Best Local Similarity 86.0%; Score 37; DB 2; Length 123;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
|||||
Db 8 GGGLVQSG 15

RESULT 4

D30539

Ig heavy chain V region (224, 2B11) - mouse (fragment)
C:Species: Mus musculus (house mouse)

C:Date: 10-Feb-1989 #sequence_revision 10-Feb-1989 #text_change 16-Aug-1996
C:Accession: D30539

R:Clatlin, J.L.; Berry, J.

J. Immunol. 141, 4012-4019, 1988

A:Title: Genetics of the phosphocholine-specific antibody response to Streptococcus pneumoniae A:Reference number: A30534; MUID:89035545

A:Accession: D30539

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation A:Molecule type: mRNA

A:Residues: 1-123 <CLA>

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:15-100/Domain: Immunoglobulin homology <IMM>

Query Match
Best Local Similarity 86.0%; Score 37; DB 2; Length 123;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
|||||

Db 8 GGGLVQSG 15

RESULT 5

I47184

Ig heavy chain variable VDJ region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)

C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47184

R:Sun, J.; Kacs Kovacs, I.; Brown, W.R.; Butler, J.E.

J. Immunol. 153, 5618-5627, 1994

A:Title: Expressed swine VH genes belong to a small VH gene family homologous to human A:Reference number: I47177; MUID:95081609

A:Accession: I47184

A:Status: preliminary; translated from GR/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-146 <SUN>

C:Cross-references: EMBL:U15437; MUID:9571372; PIDN:AAA67009.1; PID:9571373
F:34-116/Domain: Immunoglobulin homology <IMM>

Query Match
Best Local Similarity 86.0%; Score 37; DB 2; Length 146;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
|||||
Db 27 GGGLVQSG 34

RESULT 6

D31485

Ig heavy chain V regions (3-24 and 5-14) - mouse (fragment)
C:Species: Mus musculus (house mouse)

C:Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 30-May-1997
C:Accession: D31485

R:Bedzyk, W.D.; Johnson, L.S.; R Jordan, G.S.; Voss Jr., E.W.

J. Biol. Chem. 264, 1565-1569, 1989

A:Title: Comparison of variable region primary structures within an anti-fluorescein A:Reference number: A31485; MUID:89109167

A:Accession: D31485

A:Status: preliminary
A:Molecule type: protein

A:Residues: 1-36 <BED>

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin

Query Match
Best Local Similarity 83.7%; Score 36; DB 2; Length 36;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
|||||
Db 8 GGGLVQPG 15

RESULT 7

S33402

Ig heavy chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)

C:Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 17-Jul-1998
C:Accession: S33402; S36385

R:Kettlborough, C.A.; Saldanha, J.; Ansell, K.H.; Bendig, M.M.

Eur. J. Immunol. 23, 206-211, 1993

A:Title: Optimization of primers for cloning libraries of mouse immunoglobulin genes A:Reference number: S33391; MUID:93122092

A:Accession: S33402

A:Molecule type: mRNA

A:Residues: 1-38 <KET>

C:Cross-references: EMBL:X73019
A:Experimental source: strain BALB/c

R:Ansell, K.H.
submitted to the EMBL Data Library, April 1993
A:Reference number: S36376
A:Accession: S36385

A:Molecule type: mRNA
A:Residues: 1-30 <ANS>
A:Cross-references: EMBL:X73019
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:1-38/Domain: Immunoglobulin homology (fragment) <IMM>

Query Match 83.7%; Score 36; DB 2; Length 38;
Best Local Similarity 87.5%; Pred. No. 1.4;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
|||||
DB 6 GGGVQPG 13

RESULT 8
A27606

Ig heavy chain V region (p26, 9p2) - rabbit (fragments)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 08-Mar-1989 #sequence_revision 30-Jun-1991 #text_change 16-Aug-1996
C:Accession: A27606
R:Currier, S.J.; Gallardo, J.L.; Knight, K.L.
J. Immunol. 140, 1651-1659, 1988
A:Title: Partial molecular genetic map of the rabbit V-H chromosomal region.
A:Reference number: A27606; MUID:88154464
A:Accession: A27606

A:Molecule type: DNA
A:Residues: 1-59 <CUR>
A:Cross-references: GB:M19706
A:Note: the authors translated the codon TCC for residue 28 as Asp
C:Genetics:
A:Introns: 30/3
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin

Query Match 83.7%; Score 36; DB 2; Length 59;
Best Local Similarity 87.5%; Pred. No. 2.3;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
|||||
DB 8 GGGVQPG 15

RESULT 9
S36381

Ig heavy chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 23-Jul-1999
C:Accession: S36381; S33396
R:Ansell, K.H.
submitted to the EMBL Data Library, April 1993
A:Reference number: S36376
A:Accession: S36381
A:Molecule type: mRNA
A:Residues: 1-59 <ANS>
A:Cross-references: EMBL:X73012; NID:g295879; PIDN:CA51498.1; PID:g939941
R:Kettlborough, C.A.; Saldanha, J.; Ansell, K.H.; Bendig, M.M.
Eur. J. Immunol. 23, 206-211, 1993
A:Title: Optimization of primers for cloning libraries of mouse immunoglobulin genes
A:Reference number: S33391; MUID:93122092
A:Accession: S33396
A:Molecule type: mRNA
A:Residues: 3-43 <KET>
A:Cross-references: EMBL:X73012
A:Experimental source: strain BALB/c

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:14-59/Domain: Immunoglobulin homology (fragment) <IMM>

Query Match 83.7%; Score 36; DB 2; Length 59;
Best Local Similarity 87.5%; Pred. No. 2.3;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
|||||
DB 10 GGGVQPG 17

RESULT 10
S36382

Ig heavy chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 23-Jul-1999
C:Accession: S36382; S33397
R:Ansell, K.H.
submitted to the EMBL Data Library, April 1993
A:Reference number: S36376
A:Accession: S36382
A:Molecule type: mRNA
A:Residues: 1-60 <ANS>
A:Cross-references: EMBL:X73011; NID:g295878; PIDN:CA51497.1; PID:g939940
R:Kettlborough, C.A.; Saldanha, J.; Ansell, K.H.; Bendig, M.M.
Eur. J. Immunol. 23, 206-211, 1993
A:Title: Optimization of primers for cloning libraries of mouse immunoglobulin genes
A:Reference number: S33391; MUID:93122092
A:Accession: S33397

A:Molecule type: mRNA
A:Residues: 1-33 <KET>
A:Cross-references: EMBL:X73011
A:Experimental source: strain BALB/c
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:8-60/Domain: Immunoglobulin homology (fragment) <IMM>

Query Match 83.7%; Score 36; DB 2; Length 60;
Best Local Similarity 87.5%; Pred. No. 2.3;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
|||||
DB 1 GGGVQPG 8

RESULT 11
C34964

Ig heavy chain precursor V-III region (Ab21) - human
C:Species: Homo sapiens (man)
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-May-1997
C:Accession: C34964
R:Sanz, I.; Casali, P.; Thomas, J.W.; Nockins, A.L.; Capra, J.D.
J. Immunol. 142, 4054-4061, 1989
A:Title: Nucleotide sequences of eight human natural autoantibody V-H regions reveals
A:Reference number: A92830; MUID:89235232
A:Accession: C34964
A:Molecule type: mRNA
A:Residues: 1-82 <SAN>
A:Cross-references: GB:M26996
A:Note: the sequence shown differs from the authors' translation after residue 56
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin

Query Match 83.7%; Score 36; DB 2; Length 82;
Best Local Similarity 87.5%; Pred. No. 3.3;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVODG 8
 |||||
 Db 27 GGGLVQPG 34

RESULT 12

D25913
 Ig heavy chain V region (BFL1) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 16-Aug-1996
 C:Accession: D25913
 R:Lawler, A.M.; Lin, P.S.; Gearhart, P.J.
 Proc. Natl. Acad. Sci. U.S.A. 84, 2454-2458, 1987
 A:Title: Adult B-cell repertoire is biased toward two heavy-chain variable-region genes
 A:Reference number: A94148; MUID:87175692
 A:Accession: D25913
 A:Molecule type: DNA
 A:Residues: 1-94 <LAN>
 C:Superfamily: Immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:11-94/Domain: immunoglobulin homology <IMM>

Query Match 83.7%; Score 36; DB 2; Length 94;
 Best Local Similarity 87.5%; Pred. No. 3.8;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 GGGLVODG 8
 |||||
 Db 4 GGGLVQPG 11

RESULT 13

I67528
 CD33 antigen homolog - mouse (fragment)
 C:Species: Mus sp. (mouse)
 C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jan-2000
 C:Accession: I67528
 R:Chiles, J.A.; Lembezat, M.P.; Freitas, A.A.
 Eur. J. Immunol. 24, 1657-1664, 1994
 A:Title: Entry of B lymphocytes into the persistent cell pool in non-immunized mice is r
 A:Reference number: I53392; MUID:94298870
 A:Accession: I67528
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-94 <RES>
 A:Cross-references: GB:S71357; NID:9550040
 C:Gene: Ig VHx24b
 C:Superfamily: Immunoglobulin V region; immunoglobulin homology
 F:11-94/Domain: immunoglobulin homology <IMM>

Query Match 83.7%; Score 36; DB 2; Length 94;
 Best Local Similarity 87.5%; Pred. No. 3.8;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 GGGLVODG 8
 |||||
 Db 4 GGGLVQPG 11

RESULT 14

I67527
 CD33 antigen homolog - mouse (fragment)
 C:Species: Mus sp. (mouse)
 C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jan-2000
 C:Accession: I67527
 R:Chiles, J.A.; Lembezat, M.P.; Freitas, A.A.
 Eur. J. Immunol. 24, 1657-1664, 1994
 A:Title: Entry of B lymphocytes into the persistent cell pool in non-immunized mice is r
 A:Reference number: I53392; MUID:94298870
 A:Accession: I67527

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-95 <RES>
 A:Cross-references: GB:S71351; NID:9550039
 C:Gene: Ig VHx24a
 C:Superfamily: Immunoglobulin V region; immunoglobulin homology
 F:12-95/Domain: immunoglobulin homology <IMM>
 Query Match 83.7%; Score 36; DB 2; Length 95;
 Best Local Similarity 87.5%; Pred. No. 3.9;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 GGGLVODG 8
 |||||
 Db 5 GGGLVQPG 12

RESULT 15

PH0872
 Ig heavy chain V region (anti-DNA, III-3R) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 16-Aug-1996
 C:Accession: PH0872
 R:Manheimer-Lory, A.; Katz, J.B.; Pillinger, M.; Ghossein, C.; Smith, A.; Diamond, B.
 J. Exp. Med. 174, 1639-1652, 1991
 A:Title: Molecular characteristics of antibodies bearing an anti-DNA-associated idiot
 A:Reference number: PH0862; MUID:92078875
 A:Accession: PH0872
 A:Molecule type: DNA
 A:Residues: 1-97 <MAN>
 A:Note: a "g" was inserted at the position of between 72nd and 73rd of the DNA sequen
 C:Comment: This antibody is produced by Epstein-Barr virus-transformed B cell that de
 C:Superfamily: Immunoglobulin V region; immunoglobulin homology
 F:14-97/Domain: immunoglobulin homology <IMM>
 F:30-35/Region: complementarity-determining 1
 F:49-66/Region: complementarity-determining 2

Query Match 83.7%; Score 36; DB 2; Length 97;
 Best Local Similarity 87.5%; Pred. No. 3.9;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 GGGLVODG 8
 |||||
 Db 7 GGGLVQPG 14

Search completed: June 13, 2001, 14:23:10
 Job time: 741 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:30:36 ; Search time 51.57 Seconds
(Without alignments)
5.314 Million cell updates/sec

Title: PCT-US01-05825A-24
Perfect score: 43
Sequence: 1 GCGLVDPG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	83.7	111	1 HV35_MOUSE	P01804 mus musculu
2	36	83.7	113	1 HV27_MOUSE	P01796 mus musculu
3	36	83.7	113	1 HV28_MOUSE	P01797 mus musculu
4	36	83.7	113	1 HV29_MOUSE	P01798 mus musculu
5	36	83.7	113	1 HV30_MOUSE	P01799 mus musculu
6	36	83.7	113	1 HV31_MOUSE	P01800 mus musculu
7	36	83.7	113	1 HV34_MOUSE	P01803 mus musculu
8	36	83.7	115	1 HV32_MOUSE	P01801 mus musculu
9	36	83.7	115	1 HV33_MOUSE	P01802 mus musculu
10	36	83.7	115	1 HV3D_HUMAN	P01765 homo sapien
11	36	83.7	116	1 HV36_MOUSE	P01806 mus musculu
12	36	83.7	116	1 HV3R_HUMAN	P01779 homo sapien
13	36	83.7	117	1 HV17_MOUSE	P01786 mus musculu
14	36	83.7	117	1 HV3C_HUMAN	P01764 homo sapien
15	36	83.7	117	1 HV3O_HUMAN	P01776 homo sapien
16	36	83.7	117	1 HV41_MOUSE	P01811 mus musculu
17	36	83.7	117	1 HV54_MOUSE	P18525 mus musculu
18	36	83.7	118	1 HV3S_MOUSE	P01809 mus musculu
19	36	83.7	118	1 HV3V_HUMAN	P80419 homo sapien
20	36	83.7	119	1 HV37_MOUSE	P01807 mus musculu
21	36	83.7	119	1 HV38_MOUSE	P01808 mus musculu
22	36	83.7	119	1 HV3K_HUMAN	P01774 homo sapien
23	36	83.7	119	1 HV3E_HUMAN	P01775 homo sapien
24	36	83.7	119	1 HV3P_HUMAN	P01777 homo sapien
25	36	83.7	119	1 HV4O_MOUSE	P01810 mus musculu
26	36	83.7	120	1 HV3E_HUMAN	P01766 homo sapien
27	36	83.7	122	1 HV2O_MOUSE	P01789 mus musculu
28	36	83.7	122	1 HV2I_MOUSE	P01790 mus musculu
29	36	83.7	123	1 HV18_MOUSE	P01787 mus musculu
30	36	83.7	123	1 HV19_MOUSE	P01788 mus musculu
31	36	83.7	123	1 HV22_MOUSE	P01791 mus musculu
32	36	83.7	123	1 HV23_MOUSE	P01792 mus musculu
33	36	83.7	123	1 HV24_MOUSE	P01793 mus musculu

ALIGNMENTS

34	36	83.7	123	1 HV25_MOUSE	P01794 mus musculu
35	36	83.7	136	1 HV16_MOUSE	P01783 mus musculu
36	36	83.7	142	1 HV01_RAT	P01805 rattus norv
37	36	83.7	144	1 HV26_MOUSE	P01795 mus musculu
38	36	83.7	528	1 YRP3_GIALA	P25203 giardia lam
39	35	81.4	115	1 HV3E_HUMAN	P01767 homo sapien
40	35	81.4	116	1 HV05_CARAU	P19181 carassius a
41	34	79.1	364	1 OTP_PARLI	P076971 paracentrot
42	34	79.1	1733	1 VNDA_PRYKA	P33485 pseudorabie
43	33	76.7	114	1 HV3B_HUMAN	P01763 homo sapien
44	33	76.7	119	1 HV3I_HUMAN	P01770 homo sapien
45	33	76.7	121	1 HV3J_HUMAN	P01771 homo sapien

RESULT 1
HV35_MOUSE STANDARD; PRT: 111 AA.
AC P01804;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION HPC76 (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8103937; PubMed=6251474;
RA Bernard O., Gough N.M.;
RT "Nucleotide sequence of immunoglobulin heavy chain joining segments
between translocated VH and mu constant regions genes".
RL Proc. Natl. Acad. Sci. U.S.A. 77:3630-3634(1980).
CC -1- MISCELLANEOUS: THE SEQUENCE OF THE FIRST 197 RESIDUES OF THE C
REGION WAS ALSO DETERMINED & DIFFERS IN ONLY 3 POSITIONS FROM THE
CC CORRESPONDING PORTION OF THE MOUSE MORE 104E MU CHAIN.
CC PIR; A02074; MMS76.
DR InterPro: IPR003006; -
DR Pfam: PF00047; 1g; 1.
KM Immunoglobulin V region.
FT NON_TER 1 1
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 12304 MW; 0EDE98EC7348056A CRC64;

Query Match 83.7%; Score 36; DB 1; Length 111;
Best Local Similarity 87.5%; Pred. No. 3.2;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GCGLVDPG 8
Db 3 GCGLVDPG 10

RESULT 2
HV27_MOUSE STANDARD; PRT: 113 AA.
AC P01786;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION A4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=78158406; PubMed=417344;
RA vrana M., Rudikoff S., Potter M.;

"Sequence variation among heavy chains from inulin-binding myeloma proteins."
 RT Proc. Natl. Acad. Sci. U.S.A. 75:1957-1961(1978).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
 CC BINDS INULIN.
 DR PIR: A93818; AVMSAB.
 DR InterPro: IPR003006;
 DR Pfam: PF00047; 1g; 1.
 KW Immunoglobulin V region.
 FT DISULFID 22 98
 FT NON_TER 113 113
 SQ SEQUENCE 113 AA; 12675 MW; 76658C12IC598285 CRC64;

Query Match 83.7%; Score 36; DB 1; Length 113;
 Best Local Similarity 87.5%; Pred. No. 3.3;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGIVQDG 8
 DB 8 GGGIVQPG 15

RESULT 3
 HV28_MOUSE STANDARD; PRT; 113 AA.
 AC P01797;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION U61.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=78158406; PubMed=417344;
 RA Vrana M., Rudikoff S., Potter M.;
 RT "Sequence variation among heavy chains from inulin-binding myeloma
 RT proteins."
 RL Proc. Natl. Acad. Sci. U.S.A. 75:1957-1961(1978).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
 CC BINDS INULIN.
 DR PIR: B93818; AVMS61.
 DR InterPro: IPR003006;
 DR Pfam: PF00047; 1g; 1.
 KW Immunoglobulin V region.
 FT DISULFID 22 98
 FT NON_TER 113 113
 SQ SEQUENCE 113 AA; 12671 MW; 76634C12IC598285 CRC64;

Query Match 83.7%; Score 36; DB 1; Length 113;
 Best Local Similarity 87.5%; Pred. No. 3.3;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGIVQDG 8
 DB 8 GGGIVQPG 15

RESULT 4
 HV29_MOUSE STANDARD; PRT; 113 AA.
 AC P01798;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION E109.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=78158406; PubMed=417344;
 RA Vrana M., Rudikoff S., Potter M.;
 RT "Sequence variation among heavy chains from inulin-binding myeloma
 RT proteins."
 RL Proc. Natl. Acad. Sci. U.S.A. 75:1957-1961(1978).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
 CC BINDS INULIN.
 DR PIR: C93818; AVMS09.
 DR InterPro: IPR003006;
 DR Pfam: PF00047; 1g; 1.
 KW Immunoglobulin V region.
 FT DISULFID 22 98
 FT NON_TER 113 113
 SQ SEQUENCE 113 AA; 12647 MW; EE50F2FE20EDB129B CRC64;

Query Match 83.7%; Score 36; DB 1; Length 113;
 Best Local Similarity 87.5%; Pred. No. 3.3;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGIVQDG 8
 DB 8 GGGIVQPG 15

RESULT 5
 HV30_MOUSE STANDARD; PRT; 113 AA.
 AC P01799;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION ABE-47N.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=77134726; PubMed=402936;
 RA Vrana M., Rudikoff S., Potter M.;
 RT "Heavy-chain variable-region sequence from an inulin-binding myeloma
 RT protein."
 RL Biochemistry 16:1170-1175(1977).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
 CC BINDS INULIN.
 DR PIR: A90400; AVMSB7.
 DR InterPro: IPR003006;
 DR Pfam: PF00047; 1g; 1.
 KW Immunoglobulin V region.
 FT DISULFID 22 98
 FT NON_TER 113 113
 SQ SEQUENCE 113 AA; 12675 MW; 76658C16C779845E CRC64;

Query Match 83.7%; Score 36; DB 1; Length 113;
 Best Local Similarity 87.5%; Pred. No. 3.3;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGIVQDG 8
 DB 8 GGGIVQPG 15

RESULT 6
 HV31_MOUSE STANDARD; PRT; 113 AA.
 AC P01800;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION T957.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=81216632; PubMed=6787122;
 RA Rudkoff S., Potter M.;
 RT "Immunoglobulin heavy chains from anti-Inulin myeloma proteins:
 RT evidence for a new heavy chain joining segment.";
 RL J. Immunol. 127:191-194(1981).
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
 CC BINDS INULIN.
 CC PIR: A92810; AVMS57.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region.
 FT DISULFID 22
 FT NON_TER 113
 SQ SEQUENCE 113 AA; 12732 MW; 26618F626B59859E CRC64;

Query Match 83.7%; Score 36; DB 1; Length 113;
 Best Local Similarity 87.5%; Pred. No. 3.3;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVODG 8
 DB 8 GGGLVODG 15

RESULT 7
 HV34_MOUSE STANDARD; PRT; 113 AA.
 AC P01803;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V REGION AMPC1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=81216632; PubMed=6787122;
 RA Rudkoff S., Potter M.;
 RT "Immunoglobulin heavy chains from anti-Inulin myeloma proteins:
 RT evidence for a new heavy chain joining segment.";
 RL J. Immunol. 127:191-194(1981).
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
 CC BINDS INULIN.
 CC PIR: A02073; HVMSAM.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region.
 FT DISULFID 22
 FT NON_TER 113
 SQ SEQUENCE 113 AA; 12691 MW; 7A6D906AAA966E9E CRC64;

Query Match 83.7%; Score 36; DB 1; Length 113;
 Best Local Similarity 87.5%; Pred. No. 3.3;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVODG 8
 DB 8 GGGLVODG 15

RESULT 8

HV32_MOUSE STANDARD; PRT; 115 AA.
 ID HV32_MOUSE
 AC P01801;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION J606.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=82099361; PubMed=6798111;
 RA Johnson N., Stankard J., Paul L., Hood L.;
 RT "The complete V domain amino acid sequences of two myeloma Inulin-
 RT binding proteins.";
 RL J. Immunol. 128:302-307(1982).
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
 CC BINDS INULIN.
 CC PIR: C92811; AVMS06.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region.
 FT DISULFID 22
 FT NON_TER 115
 SQ SEQUENCE 115 AA; 12810 MW; B67AD6638A121A5F CRC64;

Query Match 83.7%; Score 36; DB 1; Length 115;
 Best Local Similarity 87.5%; Pred. No. 3.4;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVODG 8
 DB 8 GGGLVODG 15

RESULT 9
 HV33_MOUSE STANDARD; PRT; 115 AA.
 ID HV33_MOUSE
 AC P01802;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION W3082.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=82099361; PubMed=6798111;
 RA Johnson N., Stankard J., Paul L., Hood L.;
 RT "The complete V domain amino acid sequences of two myeloma Inulin-
 RT binding proteins.";
 RL J. Immunol. 128:302-307(1982).
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
 CC BINDS INULIN.
 CC PIR: D92811; AVMS82.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region.
 FT DISULFID 22
 FT NON_TER 115
 SQ SEQUENCE 115 AA; 12887 MW; 9B4517648C121C5A CRC64;

Query Match 83.7%; Score 36; DB 1; Length 115;
 Best Local Similarity 87.5%; Pred. No. 3.4;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVODG 8

DB 8 GGGLVQPG 15

```

RESULT 10
HY3D_HUMAN STANDARD: PRT: 115 AA.
AC P01765;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION TIL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=78005528; PubMed=409716;
RA Wang A.-C., Wang I.Y., Fudenberg H.H.;
RT "Immunoglobulin structure and genetics. Identity between variable
RL J. Biol. Chem. 252:7192-7199(1977).
CC -I- MISCELLANEOUS: THE SEQUENCES OF THE V REGIONS OF THE HEAVY CHAINS
CC OF IGM AND IGG2 ISOLATED FROM A SINGLE PATIENT WITH BICLONAL
CC GAMMOPATHY ARE IDENTICAL. THEIR LIGHT CHAINS ARE APPARENTLY ALSO
CC IDENTICAL.
DR PIR: A02048; H3HUTL.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT NON_TER 115
SQ SEQUENCE 115 AA; 12356 MW; 4DCC67D179F62326 CRC64;

```

Query Match Best Local Similarity 83.7%; Score 36; DB 1; Length 115;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
DB 8 GGGLVQPG 15

```

RESULT 11
HY36_MOUSE STANDARD: PRT: 116 AA.
AC P01806;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V REGION 441 PRECURSOR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82059449; PubMed=6795591;
RA Olio R., Aufray C., Sikorav J.-L., Rougeon F.;
RT "Mouse heavy chain variable regions: nucleotide sequence of a
RL Nucleic Acids Res. 9:4099-4109(1981).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: V00774; CAA24152.1; -

```

```

DR PIR: A02076; HWS44.
DR HSSP: P01810; 2PRT.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1
FT CHAIN 19 IG HEAVY CHAIN V REGION 441.
FT NON_TER 116
SQ SEQUENCE 116 AA; 12911 MW; 6B4FF3EF9A95D9F3 CRC64;

```

Query Match Best Local Similarity 83.7%; Score 36; DB 1; Length 116;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
DB 26 GGGLVQPG 33

```

RESULT 12
HY3R_HUMAN STANDARD: PRT: 116 AA.
AC P01779;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION TUR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=74142702; PubMed=4522793;
RA Capra J.D., Kenoe J.M.;
RT "Variable region sequences of five human immunoglobulin heavy chains
RT of the VH3 subgroup: definitive identification of four heavy chain
RT hypervariable regions."
CC Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGAI MYELOMA
CC PROTEIN.
DR PIR: A02062; A1HUTU.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT NON_TER 116
SQ SEQUENCE 116 AA; 12431 MW; EB705F553A963F0C CRC64;

```

Query Match Best Local Similarity 83.7%; Score 36; DB 1; Length 116;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
DB 8 GGGLVQPG 15

```

RESULT 13
HY17_MOUSE STANDARD: PRT: 117 AA.
AC P01786;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V REGION MOPC 47A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE.

```

RX MEDLINE=80049769; PubMed=115869;
 RA Robinson E.A.; Appella E.;
 RT "Amino acid sequence of a mouse myeloma immunoglobulin heavy chain
 (MOPC 47 A) with a 100-residue deletion.";
 RL J. Biol. Chem. 254:11418-11430(1979).
 CC -1- MISCELLANEOUS: THIS ALPHA CHAIN WAS ISOLATED FROM A MYELOMA
 PROTEIN THAT CONTAINS ONE LIGHT AND ONE HEAVY CHAIN PER MOLECULE.
 CC LINKED BY A DISULFIDE BOND. IN CONTRAST, NORMAL MOUSE IGA
 MOLECULES CONTAIN TWO LIGHT AND TWO HEAVY CHAINS AND LACK A
 LIGHT-HEAVY CHAIN DISULFIDE BOND.
 CC PIR: A02069; A1MS47.
 DR HSSP: P01789; 2MCP.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 1g.1.
 DR Immunoglobulin V region.
 KW NCBI_TaxID=9606;
 SQ SEQUENCE 117 AA; 12975 MW; 0C74BE8B8154BDF4 CRC64;

Query Match 83.7%; Score 36; DB 1; Length 117;
 Best Local Similarity 87.5%; Pred. No. 3.4;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLYODG 8
 DB 8 GGGLYODG 15

RESULT 14
 ID HV3C_HUMAN STANDARD; PRT; 117 AA.
 AC P01764;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION VH26 PRECURSOR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8101090; PubMed=6450418;
 RA Mathysens G.; Rabbits T.H.;
 RT "Structure and multiplicity of genes for the human immunoglobulin
 heavy chain variable region.";
 RL Proc. Natl. Acad. Sci. U.S.A. 77:6561-6565(1980).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: J00236; AAA53516.1; -
 DR EMBL: M35415; AAA58735.1; -
 DR PIR: A02047; H3H026.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 1g.1.
 DR Immunoglobulin V region; Signal.
 KW SIGNAL
 FT CHAIN 1 19
 FT NON_TER 20 117 IG HEAVY CHAIN V-III REGION VH26.
 SQ SEQUENCE 117 AA; 12582 MW; E826733F1A3CB0F1 CRC64;

Query Match 83.7%; Score 36; DB 1; Length 117;
 Best Local Similarity 87.5%; Pred. No. 3.4;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLYODG 8

DB 27 GGGLYODG 34

RESULT 15
 ID HV3C_HUMAN STANDARD; PRT; 117 AA.
 AC P01764;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION WAS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=74142702; PubMed=4522793;
 RA Capra J.D.; Kenoe J.M.;
 RT "Variable region sequences of five human immunoglobulin heavy chains
 of the VH3 subgroup: definitive identification of four heavy chain
 hypervariable regions.";
 RL Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA
 PROTEIN.
 DR PIR: A02059; G1HWS.
 DR HSSP: P01772; 2IG2.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 1g.1.
 DR Immunoglobulin V region.
 KW NCBI_TaxID=9606;
 SQ SEQUENCE 117 AA; 13091 MW; 201DEF0E1E53D9BF CRC64;

Query Match 83.7%; Score 36; DB 1; Length 117;
 Best Local Similarity 87.5%; Pred. No. 3.4;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLYODG 8
 DB 8 GGGLYODG 15

Search completed: June 13, 2001, 14:30:36
 Job time: 526 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:29:42 ; Search time 150.43 Seconds
(without alignments)
6.233 Million cell updates/sec

Title: PCT-US01-05825A-24
Perfect score: 43
Sequence: 1 GGGLVQDG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : SPTREMBL.15:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.protist:*
12: sp.unclassified:*
13: sp.vertebrate:*
14: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being predicted, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	38	88.4	1000	2	09PCD0	09pcd0 xylella fas
2	36	83.7	15	11	09QV16	09qv16 rattus sp.
3	36	83.7	95	4	09UL86	09ul86 homo sapien
4	36	83.7	118	4	09UL91	09ul91 homo sapien
5	36	83.7	118	4	09UL72	09ul72 homo sapien
6	36	83.7	228	2	09X655	09x655 streptomyce
7	36	83.7	249	2	084994	084994 burkholderi
8	36	83.7	290	4	095848	095848 homo sapien
9	36	83.7	302	2	P96939	P96939 mycobacteri
10	36	83.7	382	2	09WWT2	09wwt2 burkholderi
11	36	83.7	382	2	09S640	09s640 burkholderi
12	36	83.7	383	2	086045	086045 burkholderi
13	36	83.7	383	2	092348	092348 burkholderi
14	36	83.7	384	2	068143	068143 burkholderi
15	36	83.7	447	2	053137	053137 rhodococcus
16	35	81.4	254	1	028311	028311 archaeoglob
17	35	81.4	489	2	09L004	09l004 streptomyce
18	35	81.4	565	11	055136	055136 mus musculu
19	35	81.4	791	3	043129	043129 aspergillus

20	34	79.1	98	10	080841	080841 arabidopsis
21	34	79.1	254	2	09R9B1	09r9b1 burkholderi
22	34	79.1	325	2	09RX23	09rx23 deinococcus
23	34	79.1	387	2	09WWT1	09wwt1 burkholderi
24	34	79.1	387	2	09S639	09s639 burkholderi
25	34	79.1	387	2	09S638	09s638 burkholderi
26	34	79.1	387	2	09R0P9	09r0p9 burkholderi
27	34	79.1	388	2	09ZFR3	09zfr3 burkholderi
28	34	79.1	388	2	P70944	P70944 burkholderi
29	34	79.1	580	2	055856	055856 synechocyst
30	34	79.1	765	2	025403	025403 helicobacte
31	34	79.1	765	2	09ZLE8	09zle8 helicobacte
32	34	79.1	1240	10	09LHD1	09lhd1 arabidopsis
33	34	79.1	1958	14	069340	069340 pseudotrib
34	33	76.7	113	4	09UL90	09ul90 homo sapien
35	33	76.7	116	4	09UL93	09ul93 homo sapien
36	33	76.7	121	4	09UL71	09ul71 homo sapien
37	33	76.7	122	4	09UL84	09ul84 homo sapien
38	33	76.7	147	4	09Y509	09y509 homo sapien
39	33	76.7	168	2	09S2Y7	09s2y7 streptomyce
40	33	76.7	243	2	09XC10	09xc10 mycobacteri
41	33	76.7	302	2	09KRV2	09krv2 vibrio chol
42	33	76.7	350	2	043986	043986 acinetobact
43	33	76.7	368	5	09VE25	09ve25 crosophilla
44	33	76.7	408	1	074052	074052 cenarchaeum
45	33	76.7	450	2	09KHW1	09khw1 rhizobium m

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	1000 AA.
09PCD0	09PCD0			
AC	09PCD0			
DT	01-OCR-2000 (TREMBLrel. 15, Created)			
DT	01-OCR-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-OCR-2000 (TREMBLrel. 15, Last annotation update)			
DE	SERINE PROTEASE.			
GN	XPI851.			
OS	xylella fastidiosa.			
OC	Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;			
OX	Xylella.			
NCBI_TaxID=2371;				
11				
SEQUENCE FROM N.A.				
RC	STRAIN-9A5C;			
RX	MEDLINE-20365717; PubMed-10910347;			
RA	Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,			
RA	Alvarenga R., Alves L.M.C., Araya J.E., Bata G.S., Baptista C.S.,			
RA	Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,			
RA	Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,			
RA	Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,			
RA	Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,			
RA	Faciocani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.R.,			
RA	Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,			
RA	Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,			
RA	Ho P.L., Hohnselt J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,			
RA	Kriegler J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,			
RA	Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,			
RA	Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,			
RA	Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,			
RA	Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,			
RA	Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,			
RA	Nani A.J., Nobrega F.G., Nunes L.R., Oliveira M.A.,			
RA	de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,			
RA	Pelajo B.R., Pereira G.A.G., Pereira H.A., Jr., Pesquero J.B.,			
RA	Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,			
RA	de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,			
RA	da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A., Jr.,			
RA	da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,			
RA	de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,			
RA	Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,			

RA Zago M.A., Zatz M., Meldanis J., Setubal J.C.:
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*."
 RL Nature 406:151-157(2000).
 DR EMBL: AE004006; AAF84657.1; -
 DR INTERPRO: IPR000169; -
 DR INTERPRO: IPR000209; -
 DR INTERPRO: IPR002865; -
 DR PFAM: PF00082; Peptidase_S8; 2.
 DR PRINTS: PRO0723; SUBTILISIN.
 DR PRINTS: PRO1217; PRICHTEXTENS.
 DR PROSITE: PS00136; SUBTILASE ASP; UNKNOWN_1.
 DR PROSITE: PS00138; SUBTILASE SER; UNKNOWN_1.
 DR PROSITE: PS00639; THIOLESTERASE_HIS; UNKNOWN_1.
 SQ SEQUENCE 1000 AA; 105429 MW; 113491226F8DE26C CRC64;

Query Match 88.4%; Score 38; DB 2; Length 1000;
 Best Local Similarity 75.0%; Pred. No. 75;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
 |||||
 Db 464 GGGLVQDG 471

RESULT 2
 ID 090V16 PRELIMINARY; PRT; 15 AA.
 AC 090V16;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE PROLACTIN-BINDING PROTEIN (FRAGMENT).
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RX NCBI_TaxID=10118;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE=95094032; PubMed=8000909;
 RA Cohen H., Cohen O., Gagnon J.;
 RT "Serum prolactin-binding protein (PRL-BP) of human and rat are
 RL identified as 19g.";
 RL C. R. Acad. Sci., III, Sci. Vie 317:293-298(1994).
 DR HSP; P01789; IMCP.
 SQ SEQUENCE 15 AA; 1469 MW; 35BD2512FF3FA369 CRC64;

Query Match 83.7%; Score 36; DB 11; Length 15;
 Best Local Similarity 87.5%; Pred. No. 1.5;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
 |||||
 Db 8 GGGLVQDG 15

RESULT 3
 ID 09ULB6 PRELIMINARY; PRT; 95 AA.
 AC 09ULB6;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE IMMUNOGLOBULIN HEAVY CHAIN (FRAGMENT).
 GN VH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tange Y., Kayano H.;

RT "Human VH gene sequence."
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB035268; BAA87067.1; -
 DR HSSP; P01772; 2FB4.
 DR INTERPRO: IPR003006; -
 DR PFAM: PF00047; 1g; 1.
 FT NON_TER 1
 FT NON_TER 95
 SQ SEQUENCE 95 AA; 10527 MW; 90A8C6D16D22574A CRC64;

Query Match 83.7%; Score 36; DB 4; Length 95;
 Best Local Similarity 87.5%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
 |||||
 Db 7 GGGLVQDG 14

RESULT 4
 ID 09UL91 PRELIMINARY; PRT; 118 AA.
 AC 09UL91;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kalls N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL: AF035023; AAD56259.1; -
 DR INTERPRO: IPR003006; -
 DR PFAM: PF00047; 1g; 1.
 FT NON_TER 1
 FT NON_TER 118
 SQ SEQUENCE 118 AA; 12843 MW; D0633949F2AC149D CRC64;

Query Match 83.7%; Score 36; DB 4; Length 118;
 Best Local Similarity 87.5%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
 |||||
 Db 8 GGGLVQDG 15

RESULT 5
 ID 09UL72 PRELIMINARY; PRT; 118 AA.
 AC 09UL72;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kalls N.N., Berney S.M.,

RA Young D.C.:
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 RL fetus."
 CLIN. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL: AF035042; AAD56278.1; -
 DR HSSP: P01772; 2F94.
 DR INTERPRO: IPR003006; -
 DR PFM: PFM0047; 19; 1.
 FT NON_TER 1 1
 FT NON_TER 118 118
 SQ SEQUENCE 118 AA; 12872 MW; BAD1A594AB2D5CCA CRC64;

Query Match 83.7%; Score 36; DB 4; Length 118;
 Best Local Similarity 87.5%; Pred. No. 16;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
 |||||
 DB 8 GGGLVQPG 15

RESULT 6
 ID 09X655 PRELIMINARY; PRT: 228 AA.
 AC 09X655:
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE NAPI KINASE.
 GN NAPI.
 OS Streptomyces collinus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 NCBI_TaxID=42684;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TU 1892;
 RX MEDLINE=99203506; PubMed=10103039;
 RA Chen S., von Bamberg D., Hale V., Breuer M., Hardt B., Mller R.,
 RA Floss H.G., Reynolds K.A., Leistner E.;
 RT "Biosynthesis of ansatrienin (mycolactenin) and naphthomycin.
 RT Identification and analysis of two separate biosynthetic gene clusters
 RT in Streptomyces collinus Tu 1892."
 RL Eur. J. Biochem. 261:98-107(1999).
 DR EMBL: AF131877; AAD31831.1; -
 DR INTERPRO: IPR000600; -
 DR PFM: PFM0480; ROK; 1.
 KW kinase.
 SQ SEQUENCE 228 AA; 22965 MW; 9E1178D1831EFE00 CRC64;

Query Match 83.7%; Score 36; DB 2; Length 228;
 Best Local Similarity 87.5%; Pred. No. 33;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
 |||||
 DB 70 GGGLVLDG 77

RESULT 7
 ID 084994 PRELIMINARY; PRT: 249 AA.
 AC 084994:
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE FLAELLIN (FRAGMENT).
 OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
 OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
 OC Burkholderia.
 NCBI_TaxID=28450;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-E936;
 RX MEDLINE=99093016; PubMed=9877189;
 RA Winstanley C., Hales B.A., Corhill J.E., Gallagher M.J., Hart C.A.;
 RT "Flagellin gene variation between clinical and environmental isolates
 RT of Burkholderia pseudomallei contrasts with the invariance among
 RT clinical isolates."
 RL J. Med. Microbiol. 47:689-694(1998).
 DR EMBL: AF030241; AAC27443.1; -
 FT NON_TER 1 1
 FT NON_TER 249 249
 SQ SEQUENCE 249 AA; 24915 MW; EFFAC8CA527174C3 CRC64;

Query Match 83.7%; Score 36; DB 2; Length 249;
 Best Local Similarity 87.5%; Pred. No. 37;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
 |||||
 DB 74 GGGLVQTG 81

RESULT 8
 ID 095848 PRELIMINARY; PRT: 290 AA.
 AC 095848:
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE HYPOTHERICAL 31.5 KDA PROTEIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Eukaryota; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Deng Y., Madan A., Banta A., Friedman C., Trask B., Hood L., Li L.;
 RT "Cloning, chromosomal localization and sequence analysis of human
 RT Jagged2 gene."
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF111170; AAD15563.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 290 AA; 31520 MW; 835D48C3460771EC CRC64;

Query Match 83.7%; Score 36; DB 4; Length 290;
 Best Local Similarity 75.0%; Pred. No. 44;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGLVQDG 8
 |||||
 DB 237 GGGLVEEG 244

RESULT 9
 ID P96939 PRELIMINARY; PRT: 302 AA.
 AC P96939:
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE HYPOTHERICAL 29.3 KDA PROTEIN.
 GN RV0650 OR MTCY20H10.31.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 Rutter S., Seeger K., Skellon S., Squares S., Squires R., Sulston J.E.,
 Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence.";
 RL Nature 393:537-544(1998).
 DR EMBL: Z92772; CAB07107.1; -
 DR TUBERCULIST: RV0550; -
 DR INTERPRO: IPR000600; -
 DR PFAM: PF00480; ROK; 1;
 DR PROSITE: PS01125; ROK; UNKNOWN_1;
 KW Hypothetical protein.
 SQ SEQUENCE 302 AA; 29269 MW; 6E4F45D4009150FE CRC64;

Query Match
 Best Local Similarity 83.7%; Score 36; DB 2; Length 302;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGVVDG 8
 Db 139 GGGVVDG 146

RESULT 10
 O9MMT2
 ID O9MMT2 PRELIMINARY; PRT; 382 AA.
 AC O9MMT2;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
 DE FLAGELLIN (FRAGMENT).
 OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
 OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
 OC Burkholderia.
 NCBI_TaxID=28450;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=E 276;
 RA Wajnanarogana S., Tungpradabkul S., Panyim S.;
 RT "Variation of flagellin genes between virulent and non-virulent
 strains of Burkholderia pseudomallei.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF078153; AAD24680.1; -
 DR INTERPRO: IPR001492; -
 DR INTERPRO: IPR001029; -
 DR PFAM: PF00669; Flagellin_N; 1;
 DR PFAM: PF00700; Flagellin_C; 1;
 DR PRINTS: PR00207; FLAGELLIN.
 FT NON_TER 382
 SQ SEQUENCE 382 AA; 38690 MW; 681E54BF7F41F3FC CRC64;

Query Match
 Best Local Similarity 83.7%; Score 36; DB 2; Length 382;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGVVDG 8
 Db 170 GGGVVDG 177

RESULT 11
 O9S640
 ID O9S640 PRELIMINARY; PRT; 382 AA.
 AC O9S640;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)

DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
 DE FLAGELLIN (FRAGMENT).
 OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
 OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
 OC Burkholderia.
 NCBI_TaxID=28450;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=E 257;
 RA Wajnanarogana S., Tungpradabkul S., Panyim S.;
 RT "Variation of flagellin genes between virulent and non-virulent
 strains of Burkholderia pseudomallei.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF078153; AAD24678.1; -
 DR INTERPRO: IPR001492; -
 DR INTERPRO: IPR001029; -
 DR PFAM: PF00669; Flagellin_N; 1;
 DR PFAM: PF00700; Flagellin_C; 1;
 DR PRINTS: PR00207; FLAGELLIN.
 FT NON_TER 382
 SQ SEQUENCE 382 AA; 38681 MW; F681308F6903991F CRC64;

Query Match
 Best Local Similarity 83.7%; Score 36; DB 2; Length 382;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGVVDG 8
 Db 170 GGGVVDG 177

RESULT 12
 O86045
 ID O86045 PRELIMINARY; PRT; 383 AA.
 AC O86045;
 DT 01-NOV-1998 (TREMblrel. 08, Created)
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
 DE FLAGELLIN.
 GN FLIC.
 OS Burkholderia thailandensis.
 OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
 OC Burkholderia.
 NCBI_TaxID=57975;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=E264;
 RA Brett P.J., Woods D.E.;
 RT "Vaccine Strategies for Immunoprophylaxis Against Melioidosis.";
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF081500; AAC31966.1; -
 DR INTERPRO: IPR001029; -
 DR INTERPRO: IPR001492; -
 DR PFAM: PF00669; Flagellin_N; 1;
 DR PFAM: PF00700; Flagellin_C; 1;
 DR PRINTS: PR00207; FLAGELLIN.
 DR PRODOM: PD000316; -; 1;
 SQ SEQUENCE 383 AA; 38809 MW; 64D681308F690399 CRC64;

Query Match
 Best Local Similarity 83.7%; Score 36; DB 2; Length 383;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGVVDG 8
 Db 170 GGGVVDG 177

RESULT 13
 O9Z3A8
 ID O9Z3A8 PRELIMINARY; PRT; 383 AA.

AC 0923A8:
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DE 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE FLAGELLIN (FRAGMENT).
OS Burkholderia cepacia (Pseudomonas cepacia).
OC Bacteria: Proteobacteria: beta subdivision; Burkholderia group;
OC Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DMS 3027, DMS 2555;
RA Tunpgradabkul S., Wajanasarana S., Tunpboonsak S., Panyim S.;
RT "PCR-restriction endonuclease of the flagellin sequences for
RT identification of Burkholderia pseudomallei and Burkholderia cepacia
RT from clinical isolates.";
RL MOL. Cell. Probes 0:0-0(1999).
DR EMBL: AF080260; AAD12054.1; -;
DR EMBL: AF080259; AAD12053.1; -;
DR INTERPRO: IPR001029; -;
DR INTERPRO: IPR001492; -;
DR PFAM: PF00669; Flagellin_N; 1.
DR PFAM: PF00700; Flagellin_C; 1.
DR PRINTS: PR00207; FLAGELLIN.
FT NON-TER 383
SQ SEQUENCE 383 AA; 38657 MW; 1B53B946EA01070E CRC64;

Query Match 83.7%; Score 36; DB 2; Length 383;
Best Local Similarity 87.5%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVODG 8
|||||
DB 170 GGGVOKG 177

RESULT 14
O68143 PRELIMINARY; PRT; 384 AA.
AC O68143:
DT 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
DE TYPE II FLAGELLIN.
GN FLIC.
OS Burkholderia cepacia (Pseudomonas cepacia).
OC Bacteria: Proteobacteria: beta subdivision; Burkholderia group;
OC Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E243;
RX MEDLINE=98155136; PubMed=9495748;
RA Hales B.A., Morgan J.A.W., Hart C.A., Winstanley C.;
RT "Variation in flagellin genes and proteins of Burkholderia cepacia.";
RL J. Bacteriol. 180:1110-1118(1998).
DR EMBL: AF011370; AAC38199.1; -;
DR INTERPRO: IPR001029; -;
DR INTERPRO: IPR001492; -;
DR PFAM: PF00669; Flagellin_N; 1.
DR PFAM: PF00700; Flagellin_C; 1.
DR PRINTS: PR00207; FLAGELLIN.
DR PRODOM: PD000316; -; 1.
SQ SEQUENCE 384 AA; 38671 MW; 960DBA015CC18A70 CRC64;

Query Match 83.7%; Score 36; DB 2; Length 384;
Best Local Similarity 87.5%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVODG 8
|||||

DB 170 GGGVOKG 177

RESULT 15
O53137 PRELIMINARY; PRT; 447 AA.
AC O53137:
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
DE COBL.
GN Rhodococcus sp.
OS Bacteria: Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.
OX NCBI_TaxID=1831;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N186/21;
RX MEDLINE=94259308; PubMed=8200543;
RA de Koc R., Nagy I., Schoofs G., Vanderleyden J.;
RT "Sequences of the cobalamin biosynthetic genes cobK, cobL and cobM
RT from Rhodococcus sp. N186/21.";
RL Gene 143:91-93(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=N186/21;
RX MEDLINE=95138028; PubMed=7836301;
RA Nagy I., Schoofs G., Compennolle F., Proost P., Vanderleyden J.,
RA de Koc R.;
RT "Degradation of the thiocarbamate herbicide EPRC (S-ethyl
RT dipropylcarbamothioate) and biosurfacting by Rhodococcus sp. strain
RT N186/21 involve an inducible cytochrome P-450 system and aldehyde
RT dehydrogenase.";
RL J. Bacteriol. 177:676-687(1995).
DR EMBL: L21196; AAC37129.1; -;
DR INTERPRO: IPR000051; -;
SQ SEQUENCE 447 AA; 47253 MW; 1194156368922DE0 CRC64;

Query Match 83.7%; Score 36; DB 2; Length 447;
Best Local Similarity 75.0%; Pred. No. 71;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGLVODG 8
|||||
DB 364 GGGVODG 371

Search completed: June 13, 2001, 14:29:42
Job time: 543 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:27:06 ; Search time 78.71 Seconds
(without alignments)
1.953 Million cell updates/sec

Title: PCT-US01-05825A-24

Perfect score: 43

Sequence: 1 GGGLVODG 8

Scoring table:

BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	86.0	119	2	US-08-318-157B-2
2	36	83.7	15	1	US-08-331-398A-66
3	36	83.7	15	2	US-08-652-558-42
4	36	83.7	15	2	US-08-331-397B-66
5	36	83.7	15	2	US-08-759-804A-65
6	36	83.7	18	1	US-08-331-398A-55
7	36	83.7	18	1	US-08-401-908-14
8	36	83.7	18	2	US-08-331-397B-55
9	36	83.7	18	2	US-08-759-804A-54
10	36	83.7	20	2	US-08-859-921A-2
11	36	83.7	20	2	US-08-053-451B-113
12	36	83.7	26	1	US-08-471-780C-80
13	36	83.7	26	1	US-08-467-282B-80
14	36	83.7	26	2	US-08-471-282A-80
15	36	83.7	26	2	US-08-466-710C-80
16	36	83.7	26	3	US-08-468-739C-80
17	36	83.7	29	1	US-08-471-780C-119
18	36	83.7	29	1	US-08-467-282B-119
19	36	83.7	29	2	US-08-471-282A-119
20	36	83.7	29	2	US-08-466-710C-119
21	36	83.7	29	3	US-08-468-739C-119
22	36	83.7	30	1	US-07-988-925-7
23	36	83.7	30	1	US-07-977-656C-75
24	36	83.7	30	1	US-08-129-930B-75
25	36	83.7	30	2	US-08-362-780-7
26	36	83.7	30	2	US-08-765-783A-92
27	36	83.7	30	2	US-08-470-139-17

28	36	83.7	67	1	US-08-162-102C-36	Sequence 36, Appl
29	36	83.7	86	2	US-08-053-451B-126	Sequence 126, App
30	36	83.7	87	1	US-08-497-312-21	Sequence 21, Appl
31	36	83.7	87	2	US-08-765-783A-105	Sequence 105, App
32	36	83.7	98	1	US-08-211-202-116	Sequence 116, App
33	36	83.7	98	1	US-07-942-245-37	Sequence 37, Appl
34	36	83.7	98	2	US-08-428-197-48	Sequence 48, Appl
35	36	83.7	98	2	US-08-665-202-31	Sequence 31, Appl
36	36	83.7	98	5	PCT-US93-10555-48	Sequence 48, Appl
37	36	83.7	108	2	US-08-428-197-4	Sequence 4, Appl1
38	36	83.7	108	5	PCT-US93-10555-4	Sequence 4, Appl1
39	36	83.7	109	1	US-07-942-245-17	Sequence 17, Appl
40	36	83.7	109	1	US-07-942-245-21	Sequence 21, Appl
41	36	83.7	109	2	US-08-428-197-3	Sequence 3, Appl1
42	36	83.7	109	5	PCT-US93-10555-3	Sequence 3, Appl1
43	36	83.7	110	1	US-08-211-202-117	Sequence 117, App
44	36	83.7	110	3	US-08-545-809A-129	Sequence 129, App
45	36	83.7	111	1	US-08-211-202-134	Sequence 134, App

ALIGNMENTS

RESULT 1
US-08-318-157B-2
Sequence 2, Application US/08318157B
Patent No. 5874340
GENERAL INFORMATION:
APPLICANT: HANSEN, Hans J.
APPLICANT: ARMOUR, Kathryn L.
TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESS: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,157B
FILING DATE: 05-OCT-1994
CLASSIFICATION: A24
ATTORNEY/AGENT INFORMATION:
NAME: Saxe, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 18733/464
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-318-157B-2

Query Match 86.0%; Score 37; DB 2; Length 119;
Best Local Similarity 87.5%; Pred. No. 8.1;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
DB 8 GGGLVODG 15

```

RESULT 2
US-08-331-398A-66
; Sequence 66, Application US/08331398A
; Patent No. 5608039
; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira
; APPLICANT: Willingham, Mark
; APPLICANT: Fitzgerald, David
; APPLICANT: Brinkmann, Ulrich
; APPLICANT: Pai, Lee
; TITLE OF INVENTION: Single Chain B3 Antibody Fusion Proteins
; TITLE OF INVENTION: and Their Uses (as amended)
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Stewart Street Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/331.398A
; FILING DATE: 28-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/767,331
; FILING DATE: 30-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/596,289
; FILING DATE: 12-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 015280-126110US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-331-398A-66

Query Match      83.7%; Score 36; DB 1; Length 15;
Best Local Similarity 87.5%; Pred. No. 1.4;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGLVQDG 8
    |||||
Db 8 GGGLVQPG 15

RESULT 3
US-08-652-558-42
; Sequence 42, Application US/08652558
; Patent No. 586155
; GENERAL INFORMATION:
; APPLICANT: LIN, AUGUSTINE YEE-THARN
; TITLE OF INVENTION: HUMANIZED ANTIBODIES AND USES
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:

ADDRESS: BANNER & WITCOFF
STREET: 75 STATE STREET, 23RD FLOOR
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,558
; FILING DATE: JUNE 6, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IB94/00387
; FILING DATE: NOVEMBER 21, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: YANKWICH, LEON R.
; REGISTRATION NUMBER: 30,237
; REFERENCE/DOCKET NUMBER: 95,497-L
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-345-9100
; TELEFAX: 617-345-9111
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-652-558-42

Query Match      83.7%; Score 36; DB 2; Length 15;
Best Local Similarity 87.5%; Pred. No. 1.4;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGLVQDG 8
    |||||
Db 8 GGGLVQPG 15

RESULT 4
US-08-331-397B-66
; Sequence 66, Application US/08331397B
; Patent No. 5981726
; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira
; APPLICANT: Benhar, Itai
; TITLE OF INVENTION: Chimeric and Mutationally Stabilized Tumor-
; TITLE OF INVENTION: Specific Antibody Fragments, Fusion Proteins, and Uses
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Stewart Street Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/331,397B
; FILING DATE: 28-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/767,331
; FILING DATE: 30-SEP-1991
```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/596,289
FILING DATE: 12-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 015280-126120US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-331-397B-66

Query Match 83.7%; Score 36; DB 2; Length 15;
Best Local Similarity 87.5%; Pred. No. 1.4;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGLVODG 8
DB 8 GGGLVODG 15

RESULT 5
US-08-759-804A-65
Sequence 65, Application US/08759804A
Patent No. 5990296
GENERAL INFORMATION:
APPLICANT: Pastan, Ira
APPLICANT: Willingham, Mark
APPLICANT: Fitzgerald, David J.
APPLICANT: Brinkmann, Ulrich
APPLICANT: Pal, Lee
TITLE OF INVENTION: Tumor-Specific Antibody Fragments,
TITLE OF INVENTION: Fusion Proteins, and Uses Thereof
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/759,804A
FILING DATE: 03-DEC-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/331,398
FILING DATE: 28-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/767,331
FILING DATE: 30-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/596,289
FILING DATE: 12-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen L.
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 015280-126140US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-759-804A-65

Query Match 83.7%; Score 36; DB 2; Length 15;
Best Local Similarity 87.5%; Pred. No. 1.4;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGLVODG 8
DB 8 GGGLVODG 15

RESULT 6
US-08-331-398A-55
Sequence 55, Application US/08331398A
Patent No. 5608039
GENERAL INFORMATION:
APPLICANT: Pastan, Ira
APPLICANT: Willingham, Mark
APPLICANT: Fitzgerald, David
APPLICANT: Brinkmann, Ulrich
APPLICANT: Pal, Lee
TITLE OF INVENTION: Single Chain B3 Antibody Fusion Proteins
TITLE OF INVENTION: and Their Uses (as amended)
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Stewart Street Plaza
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105-1492
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,398A
FILING DATE: 28-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/767,331
FILING DATE: 30-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/596,289
FILING DATE: 12-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 015280-126110US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-331-398A-55

Query Match 83.7%; Score 36; DB 1; Length 18;

Best Local Similarity 87.5%; Pred. No. 1.7;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGIVODG 8
| | | | |
Db 8 GGGIVQPG 15

RESULT 7

US-08-401-908-14
; Sequence 14, Application US/08401908
; Patent No. 5684146

GENERAL INFORMATION:

APPLICANT: Yoshitoku OKUNO et al.
TITLE OF INVENTION: A DNA WHICH CODES FOR THE VARIABLE REGION OF
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/401,908
FILING DATE: March 10, 1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:

TELEX:

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 18

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE: heavy chain of C179

Query Match

Best Local Similarity 83.7%; Score 36; DB 1; Length 18;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGIVODG 8
| | | | |
Db 8 GGGIVQPG 15

RESULT 8

US-08-331-397B-55
; Sequence 55, Application US/08331397B
; Patent No. 5981726

GENERAL INFORMATION:

APPLICANT: Pastan, Ira
APPLICANT: Benhar, Itai
TITLE OF INVENTION: Chimeric and Mutationally Stabilized Tumor-
Specific Antibody Fragments, Fusion Proteins, and Uses

TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 68

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Stewart Street Plaza
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105-1492

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/331,397B

FILING DATE: 28-OCT-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/767,331

FILING DATE: 30-SEP-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/596,289

FILING DATE: 12-OCT-1990

ATTORNEY/AGENT INFORMATION:

NAME: Hunter, Tom

REGISTRATION NUMBER: 38,498

REFERENCE/DOCKET NUMBER: 015280-12612005

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 543-9600

TELEFAX: (415) 543-5043

INFORMATION FOR SEQ ID NO: 55:

SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-331-397B-55

Query Match 83.7%; Score 36; DB 2; Length 18;
Best Local Similarity 87.5%; Pred. No. 1.7;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGIVODG 8
| | | | |
Db 8 GGGIVQPG 15

RESULT 9

US-08-759-804A-54
; Sequence 54, Application US/08759804A
; Patent No. 5990296

GENERAL INFORMATION:

APPLICANT: Pastan, Ira

APPLICANT: Willingham, Mark

APPLICANT: Fitzgerald, David J.

APPLICANT: Brinkmann, Ulrich

APPLICANT: Pai, Lee

TITLE OF INVENTION: Tumor-Specific Antibody Fragments,
Fusion Proteins, and Uses Thereof

NUMBER OF SEQUENCES: 68

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/759,804A
FILING DATE: 03-DEC-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/331,398
FILING DATE: 28-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/767,331
FILING DATE: 30-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/596,289
FILING DATE: 12-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen L.
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 015280-126140US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-759-804A-54

Query Match 83.7%; Score 36; DB 2; Length 18;
Best Local Similarity 87.5%; Pred. No. 1.7;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGLVODG 8
DB 8 GGGLVODG 15

RESULT 10
US-08-859-931A-2
Sequence 2, Application US/08859931A
Patent No. 5945510
GENERAL INFORMATION:
APPLICANT: FASANO, Alessio
TITLE OF INVENTION: SUBSTANTIALLY PURE ZONULIN, A
TITLE OF INVENTION: PHYSIOLOGICAL MODULATOR OF
TITLE OF INVENTION: MAMMALIAN TIGHT JUNCTIONS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W., Suite 800
CITY: Washington, D.C.
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,931A
FILING DATE: 21 MAY 1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: KIT, Gordon
REGISTRATION NUMBER: 30,764
REFERENCE/DOCKET NUMBER: A-6901
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060

TELEFAX: (202) 293-7860
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
US-08-859-931A-2

Query Match 83.7%; Score 36; DB 2; Length 20;
Best Local Similarity 87.5%; Pred. No. 1.9;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGLVODG 8
DB 8 GGGLVODG 15

RESULT 11
US-08-053-451B-113
Sequence 113, Application US/08053451B
Patent No. 5955584
GENERAL INFORMATION:
APPLICANT: Chen, Francis W.
APPLICANT: Dillow, Charles C.
APPLICANT: Calenoff, Emanuel
TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC
TITLE OF INVENTION: ANTIGENS, ANTIBODIES THEREOF, AND USES THEREOF
NUMBER OF SEQUENCES: 176
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/053,451B
FILING DATE: 26-APR-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Halliuh, Albert P.
REGISTRATION NUMBER: 25,227
REFERENCE/DOCKET NUMBER: 7606-033-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-854-3660
TELEFAX: 415-854-3694
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2...3
OTHER INFORMATION: /note- "Where xaa - k or q - lys or
OTHER INFORMATION: Gln"
US-08-053-451B-113

Query Match 83.7%; Score 36; DB 2; Length 20;
Best Local Similarity 87.5%; Pred. No. 1.9;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGVODG 8
|||||
Db 7 GGGVOPG 14

RESULT 12

US-08-471-780C-80
; Sequence 80, Application US/08471780C
; Patent No. 5759808
; GENERAL INFORMATION:
; APPLICANT: Casterman, Cecile
; APPLICANT: Hamers, Raymond
; TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,780C
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/106,944
; FILING DATE: 17-AUG-1993
; APPLICATION NUMBER: FR 92402326.0
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 93401310.3
; FILING DATE: 21-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E.R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 04958.0008-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4400
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Camelus dromedarius
; US-08-471-780C-80

Query Match 83.7%; Score 36; DB 1; Length 26;
Best Local Similarity 87.5%; Pred. No. 2.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGVODG 8
|||||
Db 8 GGGVOPG 15

RESULT 13

US-08-467-282B-80
; Sequence 80, Application US/08467282B
; Patent No. 5800988
; GENERAL INFORMATION:
; APPLICANT: Casterman, Cecile

; APPLICANT: Hamers, Raymond
; TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,282B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/106,944
; FILING DATE: 17-AUG-1993
; APPLICATION NUMBER: FR 92402326.0
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 93401310.3
; FILING DATE: 21-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E.R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 04958.0008-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4400
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Camelus dromedarius
; US-08-467-282B-80

Query Match 83.7%; Score 36; DB 1; Length 26;
Best Local Similarity 87.5%; Pred. No. 2.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGVODG 8
|||||
Db 8 GGGVOPG 15

RESULT 14

US-08-471-282A-80
; Sequence 80, Application US/08471282A
; Patent No. 5840853
; GENERAL INFORMATION:
; APPLICANT: Casterman, Cecile
; APPLICANT: Hamers, Raymond
; TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,282A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potler, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958,0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Camelus dromedarius
US-08-471-282A-80

Query Match 83.7%; Score 36; DB 2; Length 26;
Best Local Similarity 87.5%; Pred. No. 2.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGVVDG 8
DB 8 GGGVDPG 15

RESULT 15
US-08-466-710C-80
Sequence 80, Application US/08466710C
Patent No. 5874541
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flanagan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,710C
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potler, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958,0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Camelus dromedarius
US-08-466-710C-80

Query Match 83.7%; Score 36; DB 2; Length 26;
Best Local Similarity 87.5%; Pred. No. 2.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGVVDG 8
DB 8 GGGVDPG 15

Search completed: June 13, 2001, 14:27:06
JOB Time: 627 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:25:45 ; Search time 150.28 Seconds
(without alignments)
7.608 Million cell updates/sec

Title: PCT-US01-05825A-25

Sequence: 98
1 EVOLVESGGLVPGCSLRL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database :

A_Geneseq_0401:*

- 1: /SID6/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SID6/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SID6/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SID6/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SID6/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SID6/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SID6/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SID6/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SID6/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SID6/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SID6/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SID6/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SID6/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SID6/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SID6/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SID6/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SID6/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SID6/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SID6/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SID6/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SID6/gcgdata/geneseq/geneseq/AA2000.DAT:*
- 22: /SID6/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	98	100.0	20 20 W94487	Human adult heart
2	98	100.0	20 21 V79130	Human adult heart
3	98	100.0	30 17 R87049	Human group III he
4	98	100.0	97 21 B40112	Anti-hIL12 antibody
5	98	100.0	98 14 R34279	Human TNF binding
6	98	100.0	98 15 R52066	Heavy chain variab
7	98	100.0	98 16 R72074	DP24 VH region. H
8	98	100.0	98 19 W59614	Anti-RSV F protein
9	98	100.0	98 21 B40084	Anti-hIL12 antibody
10	98	100.0	98 21 B40085	Anti-hIL12 antibody
11	98	100.0	98 21 B40086	Anti-hIL12 antibody

12	98	100.0	98 21 B40087	Anti-hIL12 antibody
13	98	100.0	98 21 B40094	Anti-hIL12 antibody
14	98	100.0	98 21 B40095	Anti-hIL12 antibody
15	98	100.0	98 21 B40096	Anti-hIL12 antibody
16	98	100.0	98 21 B40113	Anti-hIL12 antibody
17	98	100.0	98 21 B40130	Anti-hIL12 antibody
18	98	100.0	98 21 B40134	Anti-hIL12 antibody
19	98	100.0	98 21 B40136	Anti-hIL12 antibody
20	98	100.0	98 21 B40137	Anti-hIL12 antibody
21	98	100.0	98 21 B40138	Anti-hIL12 antibody
22	98	100.0	98 21 B40139	Anti-hIL12 antibody
23	98	100.0	98 21 B40140	Anti-hIL12 antibody
24	98	100.0	98 21 Y56660	Partial peptide fr
25	98	100.0	98 22 B46022	Heavy chain sequen
26	98	100.0	100 21 B40068	Anti-hIL12 antibody
27	98	100.0	100 21 B40069	Anti-hIL12 antibody
28	98	100.0	100 21 B40070	Anti-hIL12 antibody
29	98	100.0	105 18 W41018	Anti-glutathione a
30	98	100.0	108 15 R54787	SPA-reactive IgG h
31	98	100.0	109 15 R54786	SPA-reactive IgM h
32	98	100.0	110 18 W41017	Anti-glutathione a
33	98	100.0	112 21 Y95183	Human consensus fr
34	98	100.0	113 19 W70621	Human consensus se
35	98	100.0	113 21 Y82347	Human immunoglobul
36	98	100.0	114 16 R66330	Mus musculus anti-
37	98	100.0	114 20 W95656	Mus musculus anti-
38	98	100.0	114 20 W95657	Anti-platelet glyco
39	98	100.0	115 21 Y95210	Human IgG1 subgrou
40	98	100.0	116 20 Y29452	Human IgG1 subgrou
41	98	100.0	116 21 B30312	Human IgG1 subgrou
42	98	100.0	116 21 Y77755	Human IgG1 subgrou
43	98	100.0	117 16 R66336	Human immunoglobul
44	98	100.0	117 16 R66325	Human immunoglobul
45	98	100.0	117 16 R66301	Human immunoglobul

ALIGNMENTS

RESULT 1	
ID W94487	standard; peptide: 20 AA.
AC W94487:	
XX	
DT 21-APR-1999	(first entry)
XX	
DE Human adult heart zonulin N-terminal peptide.	
XX	
KW Zonulin; mammalian tight junction; zonula occludens toxin; ZOT;	
KW Vibri cholerae; vaccine; cholera toxin; polyclonal antibody;	
XX	
OS Homo sapiens.	
XX	
PN W09852415-A1.	
XX	
PD 26-NOV-1998.	
XX	
PF 28-APR-1998:	98MO-US07636.
XX	
PR 21-MAY-1997:	97US-0859931.
XX	
(UYMA -) UNIV MARYLAND BALTIMORE.	
PA Fasano A:	
XX	
PI WPI, 1999-070123/06.	
XX	
DR New purified zonulin - which is capable of reversibly opening	
XX	
PT mammalian tight junctions, used for enhancing the delivery of agents	
PT across intestinal and nasal mucosa and blood brain barrier	

PS Claim 2; Page 45; 64pp; English.

CC The present invention describes pure zonulin which has an apparent
 CC molecular weight of 47 kD, as determined by SDS-PAGE, which is
 CC recognised by both anti-tau polyclonal antibody and by anti-zonula
 CC occludens toxin (ZOT) polyclonal antibody, and is capable of reversibly
 CC opening mammalian tight junctions. Zonulin proteins function as
 CC physiological modulators of mammalian tight junctions. They can be used
 CC for enhancing the absorption of therapeutic agents across tight
 CC junctions of intestinal and nasal mucosa and across tight junctions of
 CC the blood brain barrier. Zonulin can be used with agents such as drugs,
 CC e.g. lidocaine, adenosine, dobutamine, dopamine, alfentanil, dezocin, nalbuphine,
 CC norepinephrine, phenolamine, doxapram, alfentanil, dezocin, nalbuphine,
 CC buprenorphine, naloxone, ketorolac, midazolam, propofol, metacurine,
 CC mivacurium, succinylcholine, cytarabine, mitomycin doxorubicin,
 CC vincristine, vinblastine, methicillin, mezlocillin, piperacillin,
 CC cefotaxime, cefonicid, cefmetazole and aztreonam, a hormone e.g.
 CC testosterone, nandrolone, menotropins, insulin, urofollitropin,
 CC interferon-alpha, interferon-beta, interferon-gamma, interleukin-1
 CC (IL-1), IL-2, IL-4, IL-8, polyvalent IgG, specific IgG, IgA, or IgM.
 CC The proteins can also be used for the production of antibodies which can
 CC be used to assay for zonulin in body tissue or fluids, or in affinity-
 CC purification of zonulin. The present sequence represents an N-terminal
 CC peptide of zonulin.

CC Sequence 20 AA:

Query Match 100.0%; Score 98; DB 20; Length 20;

Best Local Similarity 100.0%; Pred. No. 2.5e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVSGGGLVPGGSLRL 20
 Db 1 evglvesggglvpggsrlrl 20

RESULT 2

ID Y79130 standard; Peptide; 20 AA.

AC Y79130;

DT 05-JUN-2000 (first entry)

DE Human adult heart zonulin N-terminal sequence.

KW Zonulin; antagonist; zonula occludens toxin receptor;

KW human; blood-brain barrier; antiinflammatory;

KW gastrointestinal inflammation; therapy.

OS Homo sapiens.

PN WO200007609-A1.

PD 17-FEB-2000.

PF 28-JUL-1999; 99WO-US16683.

PR 03-AUG-1998; 98US-0127815.

PA (UYWA-) UNIV MARYLAND BALTIMORE.

PI Fasanio A;

DR WPI: 2000-205565/18.

PT New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis

PS Example 3; Fig 6; 69pp; English.

CC The present sequence is that of the N-terminal region of adult
 CC human heart zonulin. The N-terminal sequences of human adult and
 CC fetal zonulins (see Y79130-36) were compared with Vibrio cholerae
 CC zonula occludens toxin (ZOT) to identify a common motif thought
 CC to be involved in receptor binding. Peptide antagonists (see
 CC Y79105-29) based on this motif are useful as antiinflammatory
 CC agents for treatment of gastrointestinal inflammation, and for
 CC treatment of conditions associated with breakdown of the blood-brain
 CC barrier.

CC Sequence 20 AA:

Query Match 100.0%; Score 98; DB 21; Length 20;

Best Local Similarity 100.0%; Pred. No. 2.5e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVSGGGLVPGGSLRL 20
 Db 1 evglvesggglvpggsrlrl 20

RESULT 3

ID R87049 standard; Peptide; 30 AA.

AC R87049;

DT 25-JUN-1996 (first entry)

DE Human group III heavy chain framework 1.

KW Humanised antibody; interleukin-5; IL-5; recombinant antibody;

KW antibody engineering; monoclonal antibody; Mab; 39D10; CDR;

KW complementarity determining region; heavy chain; framework;

KW eosinophilia; allergy; asthma.

OS Homo sapiens.

PN WO9535375-A1.

PD 28-DEC-1995.

PF 16-JUN-1995; 95WO-GB01411.

PR 17-JUN-1994; 94GB-0012230.

PA (CLLT) CELLTECH THERAPEUTICS LTD.

PI Atkwal DS, Bodmer MW, Emlage JS;

DR WPI: 1996-058412/06.

PT Anti-human IL-5 recombinant antibody - useful for preventing or

PT reducing eosinophilia and for treating certain allergic diseases,

PT esp. asthma

PS Example 3; Fig 4; 69pp; English.

CC Framework regions (R87049-52) of human group III (gp3) germ line
 CC antibody heavy chain showed homology to corresponding regions
 CC (R87053-56, respectively) of the rat anti-human interleukin-5
 CC monoclonal antibody 39D10 heavy chain (see R87039). This homology
 CC was utilised in the prodn. of a humanised 39D10 VH (R87058) in
 CC which rat 39D10 VH complementarity determining regions were grafted
 CC into the human gp3 framework.

CC Sequence 30 AA:

Query Match 100.0%; Score 98; DB 17; Length 30;

Best Local Similarity 100.0%; Pred. No. 3.9e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 98; DB 14; Length 98;
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVPGGSLRL 20
 Db 1 evqlvesggglvpggsrlr 20

RESULT 6
 R52066
 ID R52066 standard; Protein: 98 AA.
 AC R52066;
 XX
 DT 11-OCT-1996 (first entry)
 XX
 DE Heavy chain variable region of human PL0123 antibody.
 XX
 KW antibody; humanised; murine; human; heavy chain; light; variable;
 KW framework region; complementarity determining region; reshaping;
 KW modelling; surface residue; modify.
 XX
 OS Homo sapiens.

Key Location/Qualifiers
 Region 1..30
 /label=framework-region_1
 /note="FR 1"
 Region 31..35
 /label=complementarity-determining-region_1
 /note="CDR 1"
 Region 36..49
 /note="FR 2"
 Region 50..59
 /note="CDR 2"
 Region 60..98
 /note="FR 3"

PN EP592106-A1.
 PD 13-APR-1994.
 XX
 PF 07-SEP-1993; 93EP-0307051.
 XX
 PR 09-SEP-1992; 92US-0942245.
 XX

PA (PEDE/) PEDERSEN J T.
 XX (IMM-) IMMUNOGEN INC.
 XX

PI Guild BC, Pedersen JT, Rees AR, Roguska MA, Searle SMJ;
 XX WPI: 1994-120230/15.
 DR

PT Method of resurfacing of rodent antibodies to produce humanised
 PT antibody forms - for producing non-human antibodies with improved
 PT therapeutic efficiency by presenting human surface on V-region
 XX
 PS Example 1: Fig 4B; 230Pp; English.

CC Modification of a rodent antibody (Ab) or fragment by resurfacing in
 CC order to produce a humanised rodent Ab can be determined by calculating
 CC homology between murine and human Ab antibody surfaces. In order to test
 CC the resurfacing approach of the invention, three humanisation experiments
 CC were set up: (1) traditional loop grafting; (2) resurfacing approach using human
 CC sequences with most similar surface residues. The Ab used was the murine
 CC anti-N901 Ab (see R52066). Experiment 3 was carried out using the present
 CC sequence which represents the human PL0123 Ab light chain variable region
 CC with 74 percent homology with anti-N901 Ab. N901/PL0123 (R52066) was
 CC prep'd. by resurfacing. Sequence numbering starts at position 118 in
 CC the specification.
 XX

SO Sequence 98 AA;

Query Match 100.0%; Score 98; DB 15; Length 98;
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVPGGSLRL 20
 Db 1 evqlvesggglvpggsrlr 20

RESULT 7
 R72074
 ID R72074 standard; Protein: 98 AA.
 AC R72074;
 XX
 DT 26-SEP-1995 (first entry)
 XX
 DE DP54 VH region.
 XX

DE DP54 VH region.
 XX
 KW Graves ophthalmopathy associated immunoglobulin protein;
 KW orbital antigen; monoclonal antibody; heavy chain; H chain;
 KW variable region; autoimmunity.
 XX
 OS Homo sapiens.

Key Location/Qualifiers
 Region 41..44
 /label=CDR1
 Region 49..67
 /label=CDR2

PN WO9508336-A.
 XX
 PD 30-MAR-1995.
 XX

PF 22-SEP-1994; 94WO-US10756.
 XX

PR 22-SEP-1993; 93US-0124469.
 XX

PA (NICH-) NICHOLS INST DIAGNOSTICS.
 XX

PI McLaughlan SM, Rapoport B;
 XX WPI: 1995-139383/18.
 DR

XX N-PSDB: Q89332.
 XX

PT Graves' ophthalmopathy-associated monoclonal antibody - produced
 PT by molecular cloning of immunoglobulin genes by PCR
 XX
 PS Disclosure; Page 72; 94Pp; English.

CC L- and H-chain DNA was amplified by PCR from Graves' orbital
 CC tissue and clones encoding autoimmune-associated immunoglobulin
 CC fragments were obtained. 2/15 clones of H chain (IgG1) genes
 CC showed homology to the germline gene DP54 (O89332), which encodes
 CC the protein given in R72074. The DNA (O89333) and corresp. amino
 CC acid (R72075) sequences of the VH region of a representative clone,
 CC OF7H1.19, are provided.
 XX

SO Sequence 98 AA;

Query Match 100.0%; Score 98; DB 16; Length 98;
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVPGGSLRL 20
 Db 1 evqlvesggglvpggsrlr 20

RESULT	8
ID	W59614 standard; Protein; 98 AA.
XX	
AC	W59614;
XX	
DT	12-OCT-1998 (first entry)
XX	
DE	Anti-RSV F protein Hu19 heavy chain GL Dp58.
XX	
KW	Monoclonal antibody; human; Hu19B; engineered antibody; RSV; respiratory syncytial virus; fusion protein; infection; complementarity determining region; CDR; therapy; diagnosis.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Region 31..35
FT	/label= CDR1
FT	Region 50..66
FT	/label= CDR2
XX	
PN	W09819704-A1.
XX	
PD	14-MAY-1998.
XX	
PE	23-OCT-1997; 97WO-US19203.
XX	
PR	01-NOV-1996; 96US-0030149.
XX	
PA	(SMIK) SMITHKLINE BEECHAM CORP.
XX	
PI	Deen KC, Dillon SB, Porter TG, Sweet RW;
XX	
DR	WPI: 1998-286600/25.
PT	Monoclonal antibodies reactive with Respiratory Syncytial Virus - useful for detection, prevention and treatment of RSV infections
PS	Disclosure: Fig 2: 10pp; English.
CC	This is the heavy chain amino acid sequence of germline (GL) Dp58 antibody. The invention relates to the construction and use of human MABs specific for the fusion (F) protein of respiratory syncytial virus (RSV) to passively treat, prevent or detect RSV infection. Hu19A, Hu19B, Hu19C and Hu19D MABs are claimed. These are reshaped human antibodies comprising a heavy chain selected from 19A, 19B, 19C or 19D (see W59615-18), which are based on the GL Dp58 sequence, and a light chain selected from 19A, 19B, 19C or 19D (see W59620-21). Such engineered antibodies are neutralising; they inhibit virus growth in vitro and in vivo in animal models of RSV infection. Nucleic acids encoding the human MABs, recombinant plasmids (see V41427-33) and host cells (e.g. COS, CHO, myeloma) are provided.
SO	Sequence 98 AA;
OY	Query Match 100.0%; Score 98; DB 19; Length 98; Best Local Similarity 100.0%; Pred. No. 1,4e-07; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
DB	1 EVOLVESGGLVQPGSLRL 20 1 evqlvesggglvqpgslrl 20
RESULT	9
ID	B40084 standard; Protein; 98 AA.
XX	
AC	B40084;

Query Match	Best Local Similarity	100.08;	Score 98;	DB 21;	Length 98;
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	EVQLVESGGGLVQPGGSLRL	20		
Db	1	evqlvesggglvqpggsrlrl	20		
RESULT	10				
ID	B40085	standard; Protein: 98 AA.			

XX BA0085;
 XX
 DT 05-FEB-2001 (first entry)
 XX
 XX Anti-IL12 antibody H chain V region amino acid sequence SEQ ID 611.
 DE
 XX Human: neutralising antibody; interleukin-12; IL-12; antiinflammatory;
 KW complementarity determining region; CDR; antirheumatic; antiarthritis;
 KW antisclerotic; neuroprotective; antispasmodic; antisthmatic; cardiant;
 KW antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis.
 OS
 XX Homo sapiens.
 PN WO200056772-A1.
 XX
 XX 28-SEP-2000.
 XX
 XX 24-MAR-2000; 2000WO-US07946.
 XX
 XX 25-MAR-1999; 99US-0126603.
 PR
 XX (BADI) BASF AG.
 PA (GEMT) GENETICS INST INC.
 XX
 XX Salfield JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M,
 PI Kaymakçalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A,
 PI Velman GM, Venturini A, Warne NW, Widom A, Elvin JG, Duncan AR,
 PI Derbyshire EJ, Carmen S, Smith S, Holter TL, Du Fou SL,
 XX WPI: 2000-638250/61.
 DR
 XX
 XX New human antibody specific for human interleukin-12 (IL-12) used to
 PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
 PT disease and multiple sclerosis -
 XX
 PS Claim 75; Page 121; 377pp; English.
 XX
 CC This invention relates to a new human antibody specific for human
 CC interleukin-12 (IL-12). The invention also includes antigen binding
 CC portions that bind to IL-12. Sequences B39485-B39516 represent human
 CC anti-IL-12 antibody heavy and light chain complementarity determining
 CC region (CDR) amino acid sequences, and also includes variable region
 CC amino acid sequences. Other variable region amino acid sequences are
 CC given in B39517-B39560 and B40068-B40149. Sequences B39561-B39771
 CC represent other IL-12 CDR3 related amino acid sequences. B39772-B40063
 CC represent other CDR sequences. Light chain CDR3 consensus sequences are
 CC given in B40064-B40067. Primers used in the identification and
 CC construction of the antibodies of the invention are given in
 CC C61062-C61071. The antibody of the invention is a neutralising antibody
 CC and has antirheumatic; antiarthritis; antisclerotic; antiinflammatory;
 CC neuroprotective; antispasmodic; antisthmatic; cardiant; antiparasitic;
 CC antibacterial; and immunosuppressive activity. The antibodies or
 CC antigen-binding fragments are useful in the treatment of disorders
 CC associated with detrimental release of human IL-12, especially Crohn's
 CC disease, multiple sclerosis and rheumatoid arthritis. They can also be
 CC used in the manufacture of a pharmaceutical composition to treat human
 CC IL-12 disorders.
 CC
 XX Sequence 98 AA;
 XX

Query Match	100.0%	Score 98;	DB 21;	Length 98;
Best Local Similarity	100.0%;	Pred. No. 1,4e-07;		
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
QY	1 EVQIVESGSGIVOPGSLRL	20		
Db	1 evqivesgsgivqpgsrlrl	20		
RESULT	11			

B40086
ID B40086 standard; Protein; 98 AA.
XX AC B40086;
XX DT 05-FEB-2001 (first entry)
XX
XX Anti-IIL12 antibody H chain V region amino acid sequence SEQ ID 612.
XX
KW Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory;
KM complementarily determining region; CDR; antirheumatic; antiarthritic;
KW antisclerotic; neuroprotective; antispasmodic; antisthmatic; cardiant;
KM antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
XX WO200056772-A1.
XX
PD 28-SEP-2000.
XX
PF 24-MAR-2000; 2000WO-US07946.
XX
PR 25-MAR-1999; 99US-0126603.
XX
PA (BADI) BASF AG.
PI (GEMX) GENETICS INST INC.
XX
PI Salfeld JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M;
PI Kaymakciyan Z, Labkovsky B, Sakorafas P, Friedrich S, Myers A;
PI Veldman GM, Venturini A, Warne NW, Widom A, Elynn JG, Duncan AR,
PI Deryshire EJ, Carmen S, Smith S, Hollet TL, Du Fou SL,
XX WPI: 2000-638250/61.
XX
PT New human antibody specific for human interleukin-12 (IL-12) used to
PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
PP disease and multiple sclerosis -
XX
PS Claim 75; Page 121; 377pp; English.
XX
CC This invention relates to a new human antibody specific for human
CC interleukin-12 (IL-12). The invention also includes antigen binding
CC portions that bind to IL-12. Sequences B39485-B39516 represent human
CC anti-IL-12 antibody heavy and light chain complementarity determining
CC region (CDR) amino acid sequences, and also includes variable region
CC amino acid sequences. Other variable region amino acid sequences are
CC given in B39517-B39560 and B40068-B40149. Sequences B39561-B39771
CC represent other CDR related amino acid sequences. B39772-B40063
CC given in B40064-C40067. Primers used in the identification and
CC construction of the antibodies of the invention are given in
CC C61062-C61071. The antibody of the invention is a neutralising antibody
CC and has antineumatic; antiarthritic; antisclerotic; antiinflammatory;
CC neuroprotective; antispasmodic; antisthmatic; cardiant; antiparasitic;
CC antibacterial and immunosuppressive activity. The antibodies or
CC antigen-binding fragments are useful in the treatment of disorders
CC associated with detrimental release of human IL-12, especially Crohn's
CC disease, multiple sclerosis and rheumatoid arthritis. They can also be
CC used in the manufacture of a pharmaceutical composition to treat human
CC IL-12 disorders.
XX
XX Sequence 98 AA;

	Query Match	100.0%	Score 98;	DB 21;	Length 98;
	Best Local Similarity	100.0%	Pred. No. 1.4e-07;		
	Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	EVQLVESGGGLVQPGGSLRL	20		
Db	1	evqlvesggglvqpggsrlr	20		

RESULT 12
ID B40087 standard; Protein; 98 AA.
XX
AC B40087;
XX
DT 05-FEB-2001 (first entry)
XX
DE Anti-hIL12 antibody H chain V region amino acid sequence SEQ ID 613.
XX
KW Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory;
KW complementarity determining region; CDR; antirheumatic; antiarthritic;
KW antisclerotic; neuroprotective; antipsoriatic; antiasthmatic; cardiant;
KW antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN WO200056772-A1.
XX
PD 28-SEP-2000.
XX
PE 24-MAR-2000; 2000WO-US07946.
XX
PF 25-MAR-1999; 99US-0126603.
XX
PA (BADI) BASF AG.
PA (GENV) GENETICS INST INC.
XX
PI Saifeld JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M;
PI Kaymakalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A;
PI Veldman GM, Venturini A, Warne NW, Widom A, Elvin JG, Duncan AR;
PI Derbyshire EJ, Carmen S, Smith S, Holtet TL, Du Fou SL;
XX
DR WPI: 2000-638250/61.
XX
PT New human antibody specific for human interleukin-12 (IL-12) used to
PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
PT disease and multiple sclerosis -
XX
PS Claim 75; Page 121; 377pp; English.
XX
XX This invention relates to a new human antibody specific for human
XX interleukin-12 (IL-12). The invention also includes antigen binding
XX portions that bind to IL-12. Sequences B39485-B39516 represent human
XX anti-IL-12 antibody heavy and light chain complementarity determining
XX region (CDR) amino acid sequences, and also includes variable region
XX amino acid sequences. Other variable region amino acid sequences are
XX given in B39517-B39560 and B40068-B40149. Sequences B39561-B39771
XX represent anti-IL-12 CDR3 related amino acid sequences, B39772-B40063
XX given in B40064-B40067. Primers used in the identification and
XX construction of the antibody of the invention are given in
XX C61062-C61071. The antibody of the invention is a neutralising antibody
XX and has antirheumatic, antiarthritic, antisclerotic, antiinflammatory,
XX neuroprotective, antipsoriatic, antiasthmatic, cardiant, antiparasitic;
XX antibacterial and immunosuppressive activity. The antibodies or
XX antigen-binding fragments are useful in the treatment of disorders
XX associated with detrimental release of human IL-12, especially Crohn's
XX disease, multiple sclerosis and rheumatoid arthritis. They can also be
XX used in the manufacture of a pharmaceutical composition to treat human
XX IL-12 disorders.
XX
SQ Sequence 98 AA:
XX
Query Match 100.0%; Score 98; DB 21; Length 98;
Best Local Similarity 100.0%; Pred. No. 1,4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EVQLVESGGGLVPGGSLRL 20
|||||

DB 1 evqlvesggglvpggsllrl 20
RESULT 13
ID B40094 standard; Protein; 98 AA.
XX
AC B40094;
XX
DT 05-FEB-2001 (first entry)
XX
DE Anti-hIL12 antibody H chain V region amino acid sequence SEQ ID 620.
XX
KW Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory;
KW complementarity determining region; CDR; antirheumatic; antiarthritic;
KW antisclerotic; neuroprotective; antipsoriatic; antiasthmatic; cardiant;
KW antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN WO200056772-A1.
XX
PD 28-SEP-2000.
XX
PE 24-MAR-2000; 2000WO-US07946.
XX
PF 25-MAR-1999; 99US-0126603.
XX
PA (BADI) BASF AG.
PA (GENV) GENETICS INST INC.
XX
PI Saifeld JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M;
PI Kaymakalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A;
PI Veldman GM, Venturini A, Warne NW, Widom A, Elvin JG, Duncan AR;
PI Derbyshire EJ, Carmen S, Smith S, Holtet TL, Du Fou SL;
XX
DR WPI: 2000-638250/61.
XX
PT New human antibody specific for human interleukin-12 (IL-12) used to
PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
PT disease and multiple sclerosis -
XX
PS Claim 75; Page 121; 377pp; English.
XX
XX This invention relates to a new human antibody specific for human
XX interleukin-12 (IL-12). The invention also includes antigen binding
XX portions that bind to IL-12. Sequences B39485-B39516 represent human
XX anti-IL-12 antibody heavy and light chain complementarity determining
XX region (CDR) amino acid sequences, and also includes variable region
XX amino acid sequences. Other variable region amino acid sequences are
XX given in B39517-B39560 and B40068-B40149. Sequences B39561-B39771
XX represent anti-IL-12 CDR3 related amino acid sequences, B39772-B40063
XX given in B40064-B40067. Primers used in the identification and
XX construction of the antibodies of the invention are given in
XX C61062-C61071. The antibody of the invention is a neutralising antibody
XX and has antirheumatic, antiarthritic, antisclerotic, antiinflammatory,
XX neuroprotective, antipsoriatic, antiasthmatic, cardiant, antiparasitic;
XX antibacterial and immunosuppressive activity. The antibodies or
XX antigen-binding fragments are useful in the treatment of disorders
XX associated with detrimental release of human IL-12, especially Crohn's
XX disease, multiple sclerosis and rheumatoid arthritis. They can also be
XX used in the manufacture of a pharmaceutical composition to treat human
XX IL-12 disorders.
XX
SQ Sequence 98 AA:
XX
Query Match 100.0%; Score 98; DB 21; Length 98;
Best Local Similarity 100.0%; Pred. No. 1,4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVPGGSLRL 20
 ||||||||||||||||
 Db 1 evqlvesggglvpggsrlr 20

RESULT 14

ID B40095 standard; Protein; 98 AA.

AC B40095;

DT 05-FEB-2001 (first entry)

DE Anti-IL12 antibody H chain V region amino acid sequence SEQ ID 621.

XX Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory;
 KW complementarily determining region; CDR; antirheumatic; antiarthritic;
 KW antisclerotic; neuroprotective; antipsoriatic; antiasthmatic; cardiant;
 KW antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis.

XX Homo sapiens.

PN WO200056772-A1.

PD 28-SEP-2000.

PF 24-MAR-2000; 2000WO-US07946.

PR 25-MAR-1999; 99US-0126603.

PA (BADI) BASF AG.

PA (GEMX) GENETICS INST INC.

PI Salfield JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M;
 PI Kaymakalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A;
 PI Veldman GM, Venturini A, Warne NW, Widom A, Elyin JG, Duncan AR;
 PI Derbyshire EJ, Carmen S, Smith S, Hollet TL, Du Fou SL;
 DR WPI: 2000-638250/61.

PT New human antibody specific for human interleukin-12 (IL-12) used to
 PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
 PT disease and multiple sclerosis -

PS Claim 75; Page 121; 377pp; English.

XX This invention relates to a new human antibody specific for human
 CC interleukin-12 (IL-12). The invention also includes antigen binding
 CC portions that bind to IL-12. Sequences B39485-B39516 represent human
 CC anti-IL-12 antibody heavy and light chain complementarily determining
 CC region (CDR) amino acid sequences, and also includes variable region
 CC amino acid sequences. Other variable region amino acid sequences are
 CC given in B39517-B39560 and B40068-B40149. Sequences B39561-B39771
 CC represent anti-IL-12 CDR3 related amino acid sequences. B39772-B40063
 CC represent other CDR sequences. Light chain CDR3 consensus sequences are
 CC given in B40064-B40067. Primers used in the identification and
 CC construction of the antibodies of the invention are given in
 CC C61062-C61071. The antibody of the invention is a neutralising antibody
 CC and has antirheumatic; antiarthritic; antisclerotic; antiinflammatory;
 CC neuroprotective; antipsoriatic; antiasthmatic; cardiant; antiparasitic;
 CC antibacterial and immunosuppressive activity. The antibodies or
 CC antigen-binding fragments are useful in the treatment of disorders
 CC associated with detrimental release of human IL-12, especially Crohn's
 CC disease, multiple sclerosis and rheumatoid arthritis. They can also be
 CC used in the manufacture of a pharmaceutical composition to treat human
 CC IL-12 disorders.

XX Sequence 98 AA;

Query Match 100.0%; Score 96; DB 21; Length 98;
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 EVOLVESGGGLVPGGSLRL 20
 ||||||||||||||||
 Db 1 evqlvesggglvpggsrlr 20

RESULT 15

ID B40096 standard; Protein; 98 AA.

AC B40096;

DT 05-FEB-2001 (first entry)

DE Anti-IL12 antibody H chain V region amino acid sequence SEQ ID 622.

XX Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory;
 KW complementarily determining region; CDR; antirheumatic; antiarthritic;
 KW antisclerotic; neuroprotective; antipsoriatic; antiasthmatic; cardiant;
 KW antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis.

XX Homo sapiens.

PN WO200056772-A1.

PD 28-SEP-2000.

PF 24-MAR-2000; 2000WO-US07946.

PR 25-MAR-1999; 99US-0126603.

PA (BADI) BASF AG.

PA (GEMX) GENETICS INST INC.

PI Salfield JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M;
 PI Kaymakalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A;
 PI Veldman GM, Venturini A, Warne NW, Widom A, Elyin JG, Duncan AR;
 PI Derbyshire EJ, Carmen S, Smith S, Hollet TL, Du Fou SL;
 DR WPI: 2000-638250/61.

PT New human antibody specific for human interleukin-12 (IL-12) used to
 PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
 PT disease and multiple sclerosis -

PS Claim 75; Page 121; 377pp; English.

XX This invention relates to a new human antibody specific for human
 CC interleukin-12 (IL-12). The invention also includes antigen binding
 CC portions that bind to IL-12. Sequences B39485-B39516 represent human
 CC anti-IL-12 antibody heavy and light chain complementarily determining
 CC region (CDR) amino acid sequences, and also includes variable region
 CC amino acid sequences. Other variable region amino acid sequences are
 CC given in B39517-B39560 and B40068-B40149. Sequences B39561-B39771
 CC represent anti-IL-12 CDR3 related amino acid sequences. B39772-B40063
 CC represent other CDR sequences. Light chain CDR3 consensus sequences are
 CC given in B40064-B40067. Primers used in the identification and
 CC construction of the antibodies of the invention are given in
 CC C61062-C61071. The antibody of the invention is a neutralising antibody
 CC and has antirheumatic; antiarthritic; antisclerotic; antiinflammatory;
 CC neuroprotective; antipsoriatic; antiasthmatic; cardiant; antiparasitic;
 CC antibacterial and immunosuppressive activity. The antibodies or
 CC antigen-binding fragments are useful in the treatment of disorders
 CC associated with detrimental release of human IL-12, especially Crohn's
 CC disease, multiple sclerosis and rheumatoid arthritis. They can also be
 CC used in the manufacture of a pharmaceutical composition to treat human
 CC IL-12 disorders.

XX Sequence 98 AA;

Query Match 100.0%; Score 98; DB 21; Length 98;
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EVQLVESGGGLVQPGGSLRL 20
 |||
 Db 1 evqlvesggglvqpggsrlr 20

Search completed: June 13, 2001, 14:25:45
 Job time: 664 sec

This Page Blank (uspto)